

TESI DOCTORAL

Títol El paper de la cadena lateral en les relacions estructura-activitat dels brassinoesteroides.

Realitzada per Sr. Marc Vilaplana Polo

en el Centre Escola Tècnica Superior IQS

i en el Departament Química Orgànica i Bioquímica

Dirigida per Dra. Carme Brosa Ballesteros

Canción del pirata

Con diez cañones por banda,
viento en popa, a toda vela,
no corta el mar, sino vuela
un velero bergantín.
Bajel pirata que llaman,
por su bravura, el Temido,
en todo mar conocido
del uno al otro confín.

La luna en el mar riel,
en la lona gime el viento,
y alza en blando movimiento
olas de plata y azul;
y ve el capitán pirata,
cantando alegre en la popa,
Asia a un lado, al otro Europa,
y allá a su frente Estambul:

«Navega, velero mío,
sin temor,
que ni enemigo navío
ni tormenta, ni bonanza
tu rumbo a torcer alcanza,
ni a sujetar tu valor.

Veinte presas
hemos hecho
a despecho
del inglés,
y han rendido
sus pendones
cien naciones
a mis pies.»

*Que es mi barco mi tesoro,
que es mi dios la libertad,
mi ley, la fuerza y el viento,
mi única patria, la mar.*

«Allá muevan feroz guerra
ciegos reyes
por un palmo más de tierra;
que yo aquí tengo por mío
cuanto abasta el mar bravío,
a quien nadie impuso leyes.

Y no hay playa,
sea cualquiera,
ni bandera
de esplendor,
que no sienta
mi derecho
y dé pecho
a mi valor.»

*Que es mi barco mi tesoro,
que es mi dios la libertad,
mi ley, la fuerza y el viento,
mi única patria, la mar.*

A la voz de «¡barco viene!»
es de ver
cómo vira y se previene
a todo trapo a escapar;
que yo soy el rey del mar,
y mi furia es de temer.

En las presas
yo divido
lo cogido
por igual;
sólo quiero
por riqueza
la belleza
sin rival.

*Que es mi barco mi tesoro,
que es mi dios la libertad,
mi ley, la fuerza y el viento,
mi única patria, la mar.*

¡Sentenciado estoy a muerte!
Yo me río;
no me abandone la suerte,
y al mismo que me condena,
colgaré de alguna entena,
quizá en su propio navío.

Y si caigo,
¿qué es la vida?
Por perdida
ya la di,
cuando el yugo
del esclavo,
como un bravo,
sacudí.

*Que es mi barco mi tesoro,
que es mi dios la libertad,
mi ley, la fuerza y el viento,
mi única patria, la mar.*

Son mi música mejor
aquilones,
el estrépito y temblor
de los cables sacudidos,
del negro mar los bramidos
y el rugir de mis cañones.

Y del trueno
al son violento,
y del viento
al rebramar,
yo me duermo
sosegado,
arrullado
por el mar.

*Que es mi barco mi tesoro,
que es mi dios la libertad,
mi ley, la fuerza y el viento,
mi única patria, la mar.*

José de Espronceda

ÍNDEX

INTRODUCCIÓ	15
1. Hormones de creixement vegetal.....	17
2. Esteroides.....	18
3. Brassinoesteroides (BRs).	18
3.1. Cronologia.....	18
3.2. Distribució en el mon vegetal.	20
3.3. Diversitat estructural.....	20
3.4. Biosíntesi.	22
3.5. Acció fisiològica.....	22
3.6. Proves de camp.	24
3.7. Bioassajos.	24
3.8. Estudis qualitius de estructura-activitat (SAR).	25
4. Brassinoesteroides a l'Institut Químic de Sarrià (IQS).....	26
4.1. Síntesi de BRs naturals.	26
4.2. Síntesi de BRs no naturals.	26
4.3. Bioassaig.....	27
4.4. Estudi d'anàlegs BRs androstànics.	27
4.5. Relacions estructura-activitat quantitatives (QSAR).....	28
1. Anàlisi conformacional.....	29
2. Selecció de la conformació activa.....	29
3. Primers estudis de QSAR (QSAR clàssic).....	29
4. Segon estudi de QSAR (MEP-QSAR).....	30
5. Tercer estudi de QSAR (CoMFA).....	30
6. Estudis qualitius estructura-activitat (SAR).	31
a) Anell A.....	31
b) Anell B.....	32
c) Cadena lateral.....	33
7. Estudi computacional del KM-01.....	34
8. Estudis quantitius estructura-activitat (QSAR).....	35
9. Conformació activa basada en mapes de GRID flexibles.....	35
5. Nomenclatura dels BRs.....	37
 OBJECTIUS	 39

TEORIA I MÈTODES	43
1. Química computacional i/o modelització molecular.....	45
2. Relacions estructura-activitat.....	45
3. Mecànica molecular.....	45
4. Anàlisi conformacional.	47
5. Descriptors moleculars.....	47
5.1. GRID.....	47
5.2. Similitud molecular i Alineament.....	49

6. Tècniques estadístiques.....	50
6.1. Regressió Lineal Múltiple (Multiple Linear Regression, MLR).	50
6.2. Anàlisi de Components Principals (Principal Component Analysis, PCA).	51
6.3. Regressió de Components Principals (Principal Component Regression, PCR).	51
6.4. Mínims Quadrats Parcial (Partial Least Squares, PLS).	51
6.5. Validació dels models de regressió.	52
7. Activitat biològica.	53
7.1. Activitat a nivell general i a nivell molecular.	53
7.2. Tipus de bioassajos.	54
7.3. Activitat i SAR.	54

DISCUSSIÓ DE RESULTATS: CAPÍTOL 1 - ANDROSTANS 55

PART I: Androstans Lliures.

1. Introducció.....	57
2. Anàlisi conformacional.	59
2.1. Dinàmica molecular (DM): Simulació i anàlisi.	60
2.2. Generació dels confòrmers.	62
2.3. Anàlisi de Components Principals (PCA).	66
3. Conformació activa i predicció de l'activitat.	68
4. Conclusions.....	69

PART II: Androstans Protegits.

1. Introducció.....	70
1.1. Síntesi d'anàlegs BRs amb cadena lateral d'androstà.	70
1.2. Activitat dels anàlegs BRs amb cadena lateral d'androstà.	71
2. Anàlisi conformacional.	72
2.1. Dinàmica molecular (DM).	73
2.2. Dinàmica molecular gestionada a través d'un procés de fusió simulada (DM-SA).	75
2.3. Anàlisi de Components Principals (PCA).	81
3. Conformació activa i predicció de l'activitat.	82
4. Conclusions.....	88

DISCUSSIÓ DE RESULTATS: CAPÍTOL 2 - CONFORMACIÓ ACTIVA DELS BRs..... 89

PART I: Obtenció.

1. Introducció.....	91
1.1. Etapes en l'elecció de la conformació activa.	91
1) Anàlisi conformacional.	91
2) Selecció de la conformació activa.	92
1.2. Conformacions actives dels BRs.....	93
2. Conformació Activa SQR.	93
2.1. Anàlisi conformacional.	93
2.2. Selecció de la conformació activa.	94
1) Conformació activa de la brassinolida.	95
2) Conformació activa de la resta d'anàlegs.	98

3. Conformació Activa IZR.....	99
3.1. Introducció.	99
3.2. Anàlisi conformacional.	99
3.3. Selecció de la conformació activa.	99
• Iteració 1.	100
• Iteració 2.	100
4. Conformació Activa MVP.....	101
4.1. Introducció.	101
1) Primera contradicció: Estudi dels anàlegs androstànics.	101
2) Segona contradicció: Interseccions de mapes de GRID.	102
3) Tercera contradicció: Hidroxils dels 22R,23R-BRs.	105
4) Quarta contradicció: Cas dels 22S,23S-BRs.	107
4.2. Anàlisi conformacional.	109
4.3. Selecció de la conformació activa.	113
• Iteració 0	113
• Iteració 1	114
• Iteració 2	116
• Iteració 3	118
• Iteració / Etapa 4.....	120
4.4. Discussió dels resultats.	121
a) Hidroxils de la brassinolida.	121
b) Cas dels 22S,23S-BRs.	123
c) Interseccions de mapes de GRID.	124
5. Conformació Activa HIP.....	125
5.1. Introducció.	125
5.2. Selecció de la conformació activa.	125
5.3. Discussió de resultats.....	127
a) Cadena lateral tipus bl.	127
b) Resta de cadenes laterals tipus.	127
c) Intersecció de mapes de GRID.	128
6. Conclusions.....	129

PART II: Comparació de metodologies.

1. Introducció.....	130
2. Comparació dels anàlisi conformacionals.....	130
2.1. Anàlisi conformacional en forma d'arbre.....	131
2.2. Anàlisi conformacional aleatori amb SYBYL.....	133
2.3. Anàlisi conformacional per dinàmica molecular i fusió simulada.	136
2.4. Anàlisi conformacional aleatori amb MOE.	137
2.5. Energia dels confòrmers segons el camp de forces.	140
3. Comparació dels mètodes de selecció de la conformació activa.	140
3.1. Metodologia basada en la comparació d'índex de similitud (SQR, RMS i MVO).	140
3.2. Metodologia basada en models de GRID flexible.....	141
4. Conclusions.....	141

DISCUSSIÓ DE RESULTATS: CAPÍTOL 3 - Q S A R 143

PART I: Conformació activa i QSAR

1. Introducció.....	145
1.1. BRs dels models (<i>data set</i>) i dades d'activitat.	146
1.2. Estratègia i pretractament de les dades.	146
2. Model CONJUNT.	147
2.1. Anàlisi de Components Principals (PCA).....	147
a) Agrupació segons les conformacions actives.	147
b) Altres agrupacions.	148
2.2. Anàlisi de mínims quadrats parcials (PLS).	149
a) Paràmetres estadístics.	149
b) Interpretació dels coeficients.	150
c) Contribucions a l'activitat.	150
2.3. Conclusions del model CONJUNT.....	151
3. Model SQR.	152
3.1. Anàlisi de Components Principals (PCA).....	152
a) Agrupació segons les cadenes laterals tipus.	152
b) Altres agrupacions.	153
3.2. Anàlisi de mínims quadrats parcials (PLS).	154
a) Paràmetres estadístics.	154
b) interpretació dels coeficients.	154
c) Contribucions a l'activitat.	155
d) Predictibilitat.	155
4. Model IZR.	157
4.1. Anàlisi de Components Principals (PCA).....	157
a) Agrupació segons les cadenes laterals tipus.	157
b) Altres agrupacions.	158
4.2. Anàlisi de mínims quadrats parcials (PLS).	158
a) Paràmetres estadístics.	158
b) interpretació dels coeficients.	159
c) Contribucions a l'activitat.	159
d) Predictibilitat.	160
5. Model MVP.....	161
5.1. Anàlisi de Components Principals (PCA).....	161
a) Agrupació segons les cadenes laterals tipus.	161
b) Altres agrupacions.	163
5.2. Anàlisi de mínims quadrats parcials (PLS).	165
a) Paràmetres estadístics.	165
b) Interpretació dels coeficients.	165
c) Contribucions a l'activitat.	167
e) Predictibilitat.	167
6. Model HIP.	169
6.1. Anàlisi de Components Principals (PCA).....	169
6.2. Anàlisi de mínims quadrats parcials (PLS).	170
a) Paràmetres estadístics.	170
b) Interpretació dels coeficients.	171
c) Contribucions a l'activitat.	171
d) Predictibilitat.	172

7. Conclusions.....	174
7.1. Quantificació de la diversitat estructural.	174
7.2. Relacions estructura-activitat quantitatives segons la conformació activa.	175
7.3. Conclusions finals.....	176

PART II: Models Reduïts.

1. Introducció.....	177
2. Compostos del “data set” i “test set”	177
3. Discussió de resultats.	179
4. Conclusions.....	180
4.1. Conclusions particulars de l'estudi.	180
4.2. Conclusions generals de l'estudi.....	180

PART III: Model HOMO.

1. Introducció.....	181
2. Model HOMO.	182
2.1. BRs del “data set”.....	182
2.2. Anàlisi de components principals (PCA).	183
2.3. Anàlisi de mínims quadrats parcials (PLS).	185
a) Paràmetres estadístics.	185
b) Contribucions a l'activitat.	185
c) Predictibilitat.	185
3. Conclusions.....	186
3.1. Conclusions particulars de l'estudi.	186
3.2. Conclusions generals de l'estudi.....	186

PART IV: Model 1 µg/planta.

1. Introducció.....	187
2. Model a 1 µg/planta.	189
2.1. BRs del “data set” i resposta.	189
2.2. Reducció de compostos.....	190
2.3. Anàlisi de mínims quadrats parcials (PLS).	190
a) Paràmetres estadístics.	190
b) Interpretació dels coeficients.	191
c) Contribucions a l'activitat.	192
d) Predictibilitat.	192
3. Conclusions.....	193

DISCUSSIÓ DE RESULTATS: CAPÍTOL 4 - ACTIVITAT 195

1. Introducció.....	197
2. Test d'inclinació de la làmina d'arròs.....	197
2.1. Protocol.	197
2.2. Fonts d'error i limitacions del test.	198
2.3. Tractament estadístic de les dades.	198
1) Tractament estadístic en cada experiment.	198
2) Tractament estadístic per cada repetició de l'experiment.....	199
2.4. Corba dosi-resposta.	199

3. Error experimental.	200
3.1. Error experimental de la resposta a una dosi concreta.	201
a) <i>Error en la resposta d'un únic experiment.</i>	201
b) <i>Error en la resposta de diversos experiments.</i>	201
3.2. Error experimental del valor d'activitat.	201
4. Factors que influeixen en la corba dosi-resposta.	202
4.1. Nombre de repeticions i experiments erronis.	202
4.2. Reducció de punts experimentals.	205
a) <i>Influència en un bioassaig.</i>	205
b) <i>Influència en múltiples bioassajos.</i>	206
c) <i>Nova proposta de reducció de punts experimentals.</i>	206
d) <i>Influència de la reducció de punts experimentals en la corba dosi-resposta.</i>	208
4.3. Rectificació de les dades.	210
4.4. Resposta experimental.	211
a) <i>Mitjana de totes les respostes individuals vs. mitjana dels experiments.</i>	211
b) <i>Mitjana dels experiments vs. experiments individuals.</i>	212
4.5. Factors extrínsecs. Corbes dosi-resposta de la brassinolida i la castasterona.	213
• <i>Origen dels compostos.</i>	217
• <i>Origen i/o envelliment de les llavors.</i>	217
• <i>Aplicació de la dosi correcta.</i>	218
• <i>Experimentador.</i>	218
• <i>Factors indeterminats.</i>	218
• <i>Conclusions.</i>	219
5. Discussió de resultats.	219
5.1. Error experimental i QSAR.	219
5.2. Error experimental de la resposta a 1 µg/planta.	220
5.3. Error experimental de l'activitat (-log(dosi_{50%})).	220
6. Conclusions.	221
 CONCLUSIONS	 223
 PART EXPERIMENTAL	 229
1. Maquinari.	231
1.1. Ordinadors Personals (PC).	231
1.2. Estacions de treball (Workstations).	231
2. Programari.	231
3. Modelització Molecular.	233
3.1. Construcció dels compostos.	233
3.2. Anàlisi Conformacional.	233
• <i>Dinàmica molecular (DM):</i>	233
• <i>Dinàmica molecular gestionada a través d'un procés de fusió simulada (DM-SA) :</i>	233

3.3. Alineament molecular.	234
• Alineament molecular a través del Potencial Molecular Electrostàtic (MEP).	234
• Alineament molecular a través dels anells C i D de l'esteroides.	234
• Alineament molecular a través dels camps moleculars de GRID.	234
3.4. 3D-QSAR.....	234
• GRID	234
• GOLPE.....	235
• Predicció de l'activitat	235
• Intersecció dels mapes de GRID	236
4. Estadística.	236
4.1. Proves d'hipòtesi.	236

BIBLIOGRAFIA	237
--------------------	-----

ANNEXOS	CD
---------------	----

INTRODUCCIÓ

1. Hormones de creixement vegetal.

El creixement d'una planta ve determinat per el seu codi genètic i per factors externs com són la llum, la temperatura, l'accés a l'aigua i als nutrients, etcètera. A nivell molecular la regulació d'aquest creixement es realitza mitjançant uns compostos que s'anomenen fitohormones o hormones vegetals.¹⁻³ En general, les hormones són productes de secreció de certs òrgans del cos animal i de les plantes les quals un cop transportades per la sang o la saba i en molt baixa concentració, exciten, inhibeixen o regulen l'activitat d'altres òrgans. A diferència de les hormones animals, la biosíntesi de les hormones vegetals no està relacionada directament amb una glàndula, sinó que està gairebé present en totes les cèl·lules.

Es creu que les fitohormones poden actuar segons dos tipus de mecanismes.⁴⁻¹¹ En el primer, l'hormona es capaç de travessar la membrana cel·lular i un cop en el citoplasma s'uneix amb el seu receptor formant un complex. Aquest complex dissociat o sense dissociar, entra en el nucli i afecta a la síntesi d'ARN missatger. En el segon, l'hormona s'uneix a un receptor de membrana formant un complex en el qual es provoca un canvi conformational del receptor que desencadena una cascada de reaccions amb efectes diversos en el citoplasma.

A nivell microscòpic les fitohormones promouen la síntesi de proteïnes, faciliten el transport de productes fotosintètics, regulen el nivell d'aigua de la planta, etcètera. A nivell macroscòpic incrementen la floració, acceleren o retarden la maduració dels fruits, disminueixen la seva caiguda, etcètera. Per totes aquestes propietats les fitohormones són utilitzades en agricultura com a reguladors del creixement vegetal.

Clàssicament, les fitohormones es classifiquen en cinc grans classes: auxines, gibberel·lines, citoquinines, l'àcid abscísic i l'etilè (*Figura 1*). No obstant, hi ha altres compostos i famílies de compostos de naturalesa molt diversa que també són considerats com fitohormones.

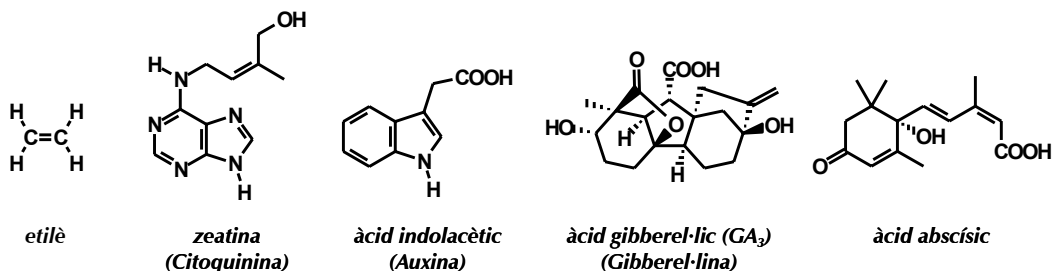


Figura 1: Els cinc grups clàssics de fitohormones.

La cronologia de les fitohormones és:

- 1901 \Rightarrow etilè:¹² estimula l'obertura de les flors, la maduració del fruit i la caiguda de les fulles. També contribueix al creixement i a la diferenciació entre l'arrel i la tija.
- 1913 \Rightarrow citoquinines:¹³ augmenten la divisió cel·lular, participen en la floració, la degradació de la clorofil·la i inhibeixen la caiguda de les fulles. Com exemples hi ha la zeatina, la quinetina i la benciladenina.
- 1926 \Rightarrow auxines:^{14,15} afavoreixen l'elongació cel·lular, contribueixen al creixement de la planta i a la maduració dels fruits. L'exemple més característic és l'àcid indolacètic (IAA).
- 1935 \Rightarrow gibberel·lines:¹⁶ indueixen la germinació de les llavors i determinen el creixement de la tija provocant l'elongació dels brots i els entrenusos. L'exemple més característic és l'àcid gibberèl·lic (GA_3).

- 1958 \Rightarrow morfactines:¹⁷ formada per diversos fluorens substituïts.
- 1963 \Rightarrow l'àcid abscísic (ABA):¹⁸ propicia la caiguda de les fulles, frena el creixement de la tija i inhibeix la germinació de les llavors.
- 1971 \Rightarrow poliamines:¹⁹ alguns exemples són la putrescina, la cadaverina i l'espermidina.
- 1979 \Rightarrow brassinoesteroides:²⁰ alguns exemples són la brassinolida i la castasterona.
- 1990 \Rightarrow jasmonats.²¹

La present tesi es centra en el camp dels **brassinoesteroides (BRs)** els quals formen part d'una extensa família de productes naturals: els esteroides.

2. Esteroides.

Els esteroides són molècules complexes que es caracteritzen per tenir l'esquelet fonamental del ciclopentanperhidrofenantrè parcial o completament hidrogenat, normalment amb grups metílics a C10 i C13 i sovint amb grups alquilics a C17. Per extensió, pot tenir lloc una o més escissions d'enllaços, expansions dels anells i/o contraccions de l'esquelet. La gran diversitat estructural que presenten està associada a una gran diversitat d'activitats biològiques de gran transcendència vital.

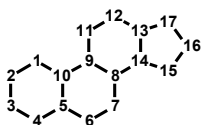


Figura 2: Esquelet de ciclopentanperhidrofenantrè i numeració IUPAC.

Dins el regne animal els esteroides abasten compostos amb funció tan diversa com els zooesterols representats pel colesterol, els àcids biliars, la vitamina D, les hormones sexuals masculines i femenines, les hormones corticoides o corticosteroides, ecdiesteroides i neurosteroides.

En el regne vegetal també presenten múltiples funcions: fitoesterols, bufanolides, brassinoesteroides, els alcaloides esteroidals o sapogenines, ecdiesteroides i cardenolides. Abans de l'aïllament de la brassinolida hi va haver una llarga discussió sobre si els esteroides realitzaven o no funcions hormonals en les pròpies plantes.²²⁻²⁴

3. Brassinoesteroides (BRs).

3.1. Cronologia.

L'any 1941 Mitchell i Whitehead del Betsville Agricultural Research Center (BARC) que pertanyia al United States Department of Agriculture (USDA) van observar que extractes del pol·len del blat de moro (*Zea mays*) tenien activitat com a promotors del creixement vegetal.²⁵ Des de llavors es va analitzar el pol·len d'unes 60 espècies, de les quals unes 20 van donar activitat biològica en el test del segon internodi de la mongeta.²⁶ L'any 1970 van publicar a *Nature* un article titulat "*Brassins – A new family of plant hormones from rape pollen*".²⁷ Tot i que els *brassins* eren principalment una barreja d'àcids grassos de glicerol²⁶ i glucosa,²⁸ semblava que la seva activitat s'havia de deure a algun altre component minoritari. L'any 1974 la USDA va iniciar un projecte per aïllar el principi actiu dels *brassins* a partir de 227 kg de pol·len de colza (*Brassica napus* L.). L'any 1979 Grove i col·laboradors van aconseguir aïllar 4 mg d'una substància cristal·lina que van anomenar brassinolida (bl,

Figura 3),²⁰ la qual van caracteritzar i identificar mitjançant dades espectroscòpiques i RX. Aquest any és considerat com l'any del descobriment dels BRs i, de fet, la majoria d'articles comencen fent referència a aquesta data i a aquest autor.

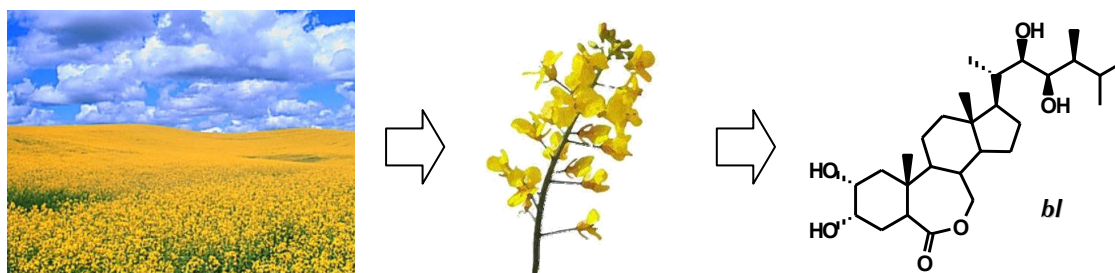


Figura 3: Camp de colza (*Brassica napus*), detall de les flors i estructura de la brassinolida (bl).

El descobriment de la brassinolida no va passar desapercebut a la Universitat de Nagoya al Japó. Des de 1966 la Universitat treballava amb uns extractes vegetals (*Dystilium* factors) que presentaven activitat similar a les auxines en el test d'inclinació de la làmina d'arròs. Aquests extractes presentaven unes propietats químiques molt similars a les descrites per la brassinolida.²⁹ L'any 1982 Yokota i col·laboradors van aïllar del castanyer japonès (*Castanea crenata*) el segon brassinoesteroide natural que van anomenar castasterona³⁰ (cs, Figura 4). De bon principi van postular que es tractava del precursor biogenètic de la brassinolida, la qual cosa van acabar demostrant l'any 1990.³¹ Prèviament, l'any 1984 van demostrar que els "*Dystilium* factors" eren una barreja de brassinoesteroides.³²



Figura 4: Arbre del castanyer japonès (*Castanea crenata*), detall de les flors i estructura de la castasterona (cs).

Des del descobriment dels BRs s'han realitzat arreu del món nombrosos treballs per identificar-los dins del regne vegetal, determinar la seva acció fisiològica, establir el seu metabolisme, sintetitzar-los en el laboratori i l'avaluar la seva resposta en proves de camp. Per aconseguir-ho s'han desenvolupat nombroses tècniques de purificació i anàlisi, diferents bioassajos, nombroses estratègies sintètiques, estudis de genètica molecular i estudis de correlació estructura-activitat. Fins l'actualitat, s'han publicat més de 1.200 articles, uns 300 articles de revisió, 5 llibres i més de 150 patents.

3.2. Distribució en el mon vegetal.

Els BRs es troben distribuïts de forma omnipresent en el regne vegetal.³³ S'han trobat en almenys 44 espècies vegetals que inclouen 37 angiospèrmies tan monocotiledònies com dicotiledònies, 5 gimnospèrmies, 1 alga i 1 molsa. Es troben en un marge de concentracions molt ampli que va des de 0,1 fins a 10^{-7} nmol/gram. Aquestes concentracions depenen molt de l'espècie, del període de desenvolupament i de la part de la planta. Les majors concentracions s'aïllen en les parts amb major activitat reproductora com el pol·len, les flors, les llavors i els petits brots.

3.3. Diversitat estructural.

Els BRs naturals tenen en comú un esquelet de 5α -colestà (Figura 5). Les seves variacions estructurals venen descrites pel tipus de funcionalitat i/o la seva orientació en tres zones ben diferenciades: L'anell A, l'anell B i la cadena lateral. S'ha arribat a aïllar i caracteritzar fins a 40 BRs i 4 conjugats els quals es poden classificar segons el nombre d'àtoms de carboni en la seva estructura.³³ No obstant, és molt més útil classificar-los segons les principals funcionalitats.

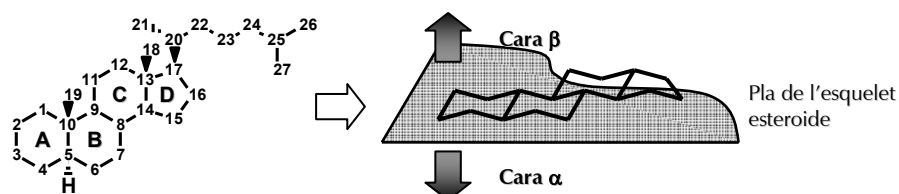


Figura 5: Esquelet de 5α -colestà. Numeració, definició dels anells i de les cares α i β .

En l'anell A la diversitat estructural és bastant gran. S'identifiquen 10 tipus de modificacions estructurals segons el nombre de grups funcionals presents, del tipus de funcionalitat, la seva posició i la seva orientació (Figura 6). La majoria de BRs naturals (19 de 44) presenten el diol $2\alpha,3\alpha$. Hi ha dos derivats on l'hidroxil en 3β està esterificat en forma de laurat i miristat.

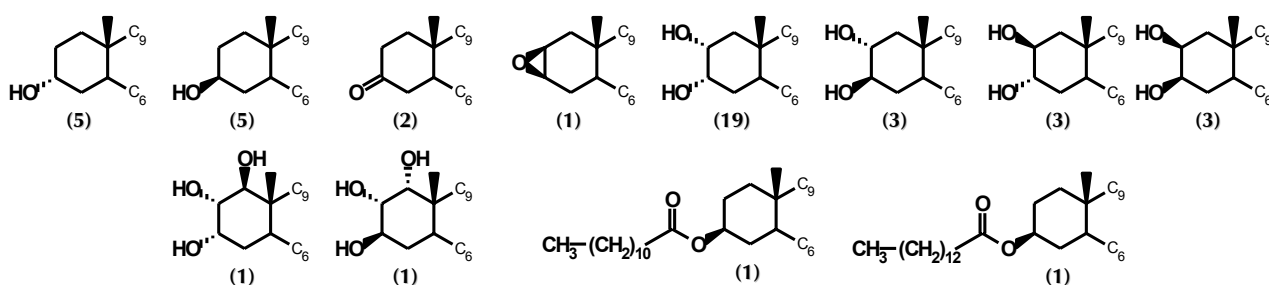


Figura 6: Diversitat estructural dels BRs naturals en l'anell A. Entre parèntesi s'indica el nombre de BRs naturals identificats amb aquesta estructura.

En l'anell B la diversitat és més reduïda ja que només s'han trobat 3 modificacions estructurals (Figura 7). L'anell B és un punt important en la diferenciació dels BRs i genèricament es parla de lactones i cetones. Tot i que els BRs més actius són lactones, la majoria de BRs naturals aïllats són cetones.

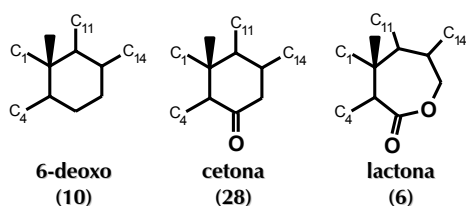


Figura 7: Diversitat estructural dels BRs naturals en l'anell B i nomenclatura. Entre parèntesi s'indica el nombre de BRs naturals identificats amb aquesta estructura.

La **cadena lateral** es va classificar inicialment en tres grups segon el nombre d'àtoms de carboni presents: C_{27} -BRs, C_{28} -BRs i C_{29} -BRs (Figura 8). Aquests grups donen pas a 8 modificacions estructurals i un derivat.

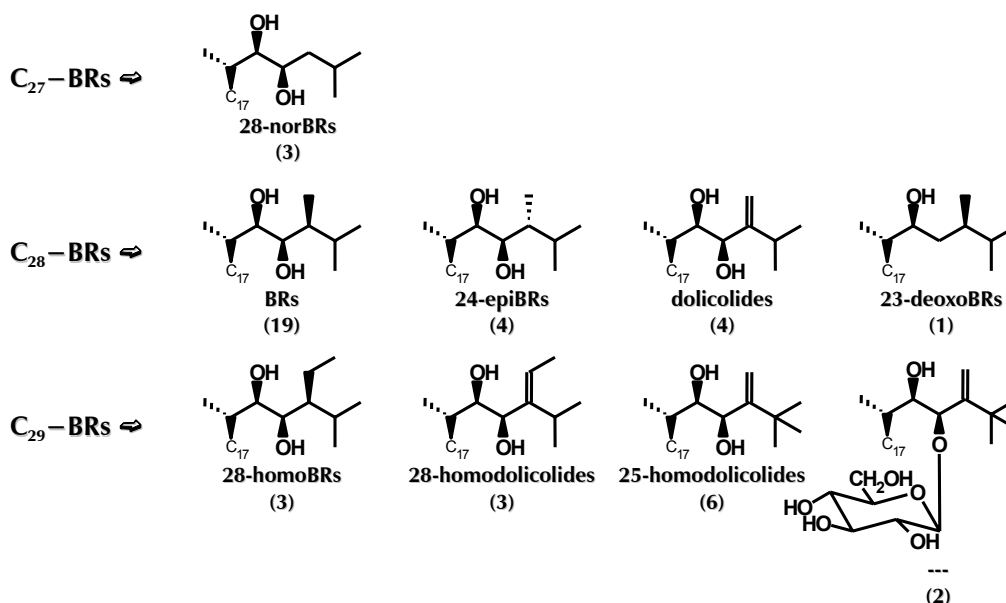


Figura 8: Diversitat estructural dels BRs naturals en la cadena lateral i nomenclatura. Entre parèntesi s'indica el nombre de BRs naturals identificats amb aquesta estructura.

Actualment la cadena lateral és el principal criteri de classificació dels BRs. Es parla de brassinolides quan la cadena lateral és la de la brassinolida, epibrassinolides (epiBRs) quan la cadena es la de la 24-epibrassinolida, homobrassinolides (homoBRs) quan la cadena es la de la 28-homobrassinolida, etcètera. Un altre punt important és la configuració del diol en C22,C23. Tot i que tots els BRs naturals aïllats presenten el diol en configuració 22*R*,23*R*, per síntesi també s'obté el diastereoisòmer 22*S*,23*S*. Això dona lloc als 22*S*,23*S*-brassinolides (Figura 9).

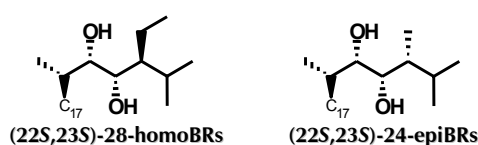


Figura 9: Exemple de dos cadenes laterals de 22*S*,23*S*-BRs i nomenclatura.

3.4. Biosíntesi.

L'any 1997 Sakurai i Fujioka van establir la ruta biogenètica dels BRs mitjançant estudis bioquímics sobre cultius de cèl·lules de *Catharanthus roseus*.^{34,35} La brassinolida es sintetitza a partir del campesterol per dos camins diferents (Figura 10): a) ruta de la prompte oxidació³⁵ (*early oxidation pathway*) i b) ruta de l'oxidació tardana³⁴ (*late oxidation pathway*). Estudis en mutants nans de la rabanisa (*Arabidopsis thaliana*) i el pèsol (*Pisum sativum*) deficients en la síntesi de BRs han confirmat que ambdós rutes biosintètiques operen en diverses espècies vegetals.

L'observació suggereix que la ruta de la prompte oxidació opera predominantment en condicions de fosc, mentre que la ruta de l'oxidació tardana té lloc predominantment en condicions de llum.^{36,37}

Degut a la complexitat de la pauta que segueixen els BRs, es creu que poden haver-hi altres camins incloent ramificacions i modificacions.³⁸

3.5. Acció fisiològica.

Els estudis dels efectes fisiològics van començar amb els *brassins* als Estats Units i els "*Distylium factors*" al Japó molt abans del descobriment de la brassinolida. A nivell cel·lular actuen augmentant la divisió cel·lular així com la mida i la elongació de la cèl·lula.³⁹ Tenen un paper important en processos de fotoregulació⁴⁰ i en el control hormonal de les plantes.⁴¹ Afavoreixen la germinació de les llavors, augmenten l'alçada de les plantes, el gruix de la tija, la inclinació de les fulles, etcètera.^{42,43} De fet, la manca de BRs condueix al nanisme.

Aplicats en proves de camp, incrementen el rendiment i la qualitat de diversos tipus de cultiu ja siguin cultius extensius com l'arròs, el blat, la soja, el blat de moro... com cultius d'alt valor afegit com bolets o plantes ornamentals. També s'han descrit que confereixen a les plantes resistència a: a) agents químics com herbicides i pesticides,^{44,45} b) factors mediambientals adversos^{46,47} com una elevada salinitat,^{48,49} exposició a temperatures extremes,^{50,51} l'escassetat de nutrients, sequera...^{52,53} i c) infeccions víriques i fongs.⁴²

Els BRs presenten un efecte sinèrgic amb altres fitohormones, especialment les auxines^{54,55} i les gibberel·lines.⁵⁶ S'ha demostrat la seva influència en l'equilibri hormonal i en la concentració de les altres fitohormones.

Per semblances estructurals dels BRs amb els ecdiesteroides, se'ls ha descrit la possible activitat antiecdiesteroida en la muda dels insectes.^{57,58}

Els BRs en tan que són constituents naturals de les plantes han estat, estan i estaran presents en la cadena tròfica dels humans. Per tant, no és probable que la seva aplicació en l'agricultura tingui efectes perjudicials per la salut humana. De fet, tots els tests per avaluar la toxicitat dels BRs en mamífers han mostrat una toxicitat nul·la o molt baixa, un efecte acumulatiu molt petit i l'absència d'activitat mutagènica.⁴² La toxicitat de la 24-epibrassinolida en rata mitjançant l'administració oral expressada com a LD₅₀ és superior a 2.000 mg/kg/dia.⁴² Aquest valor està entre els 975 mg/kg/dia de l'àcid cítric⁵⁹ i els 3.750 mg/kg/dia del clorur sòdic.⁶⁰ Tots dos són àmpliament utilitzats com additius alimentaris.

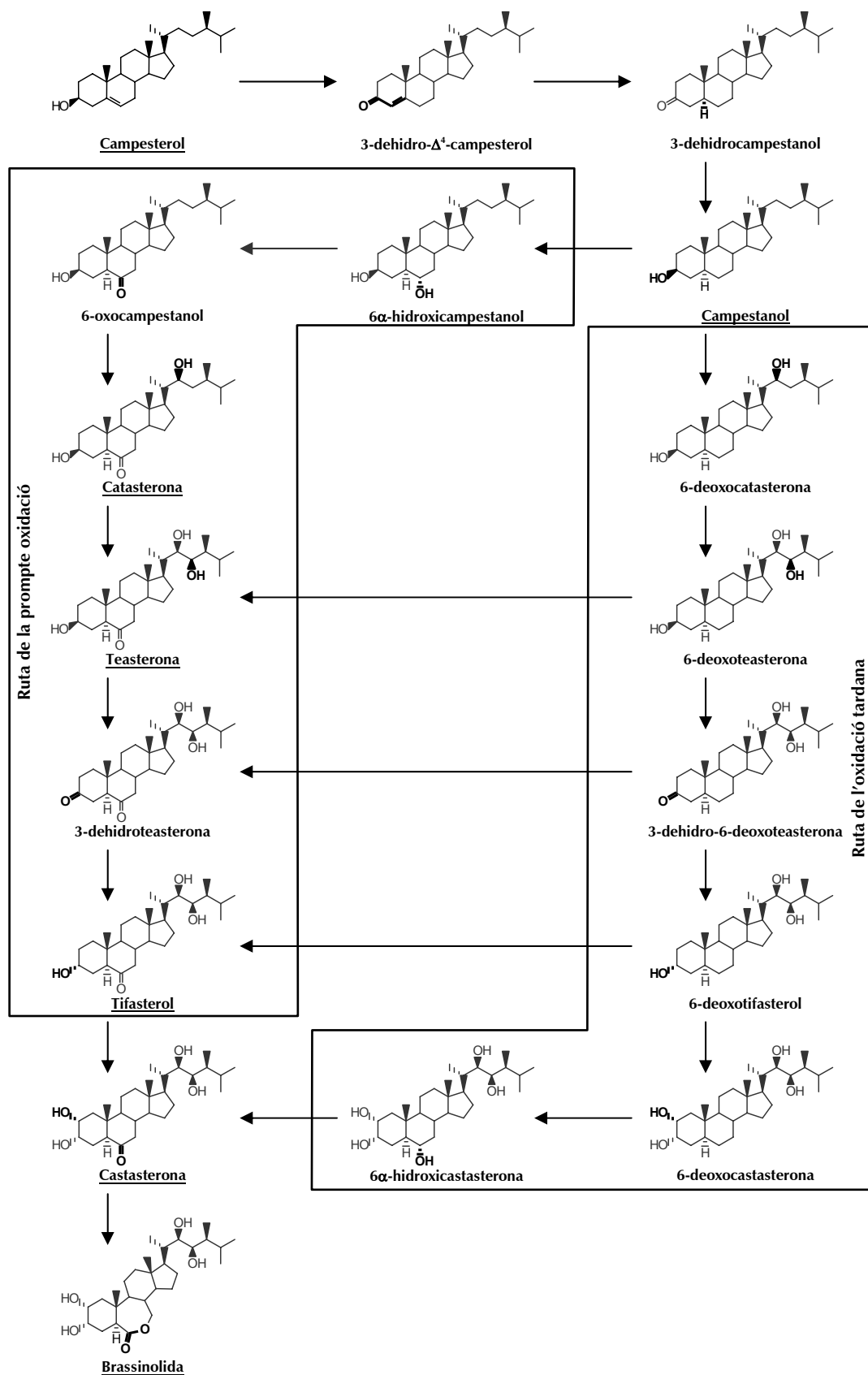


Figura 10: Ruta biosintètica dels BRs i nomenclatura utilitzada per Sakurai.³⁸

3.6. Proves de camp.

A finals de la dècada dels 80 va començar a estudiar-se l'efecte dels BRs sobre diferents tipus de cultius.^{61,62} Des d'aleshores, s'han publicat nombrosos articles i molts dels seus usos han estat patentats. Els compostos estudiats que han mostrat un major efecte en les aplicacions a camp són la (22S,23S)-28-homobrassinolida i la 24-epibrassinolida (Figura 11).⁶³

L'aplicació extensiva dels BRs topa amb dos limitacions:

- L'accessibilitat als BRs és encara baixa ja que, per una banda, la seva extracció del medi natural és inviable econòmicament per trobar-se en molt baixa quantitat i per l'altre banda, la síntesi de BRs naturals és llarga, complexa i igualment costosa a nivell econòmic.
- El ràpid metabolisme i la conseqüent inactivació dels BRs naturals.⁶⁴ Aquest fet obliga a realitzar diverses aplicacions en diferents etapes del creixement de la planta.

Per solucionar aquest problema s'han buscat anàlegs estructurals amb una elevada estabilitat en condicions de camp encara que tinguessin una baixa activitat biològica a nivell de bioassaig. L'anàleg TS303 (Figura 11) té una estabilitat i un comportament superior a la brassinolida en una gran varietat de condicions de camp.⁶⁵

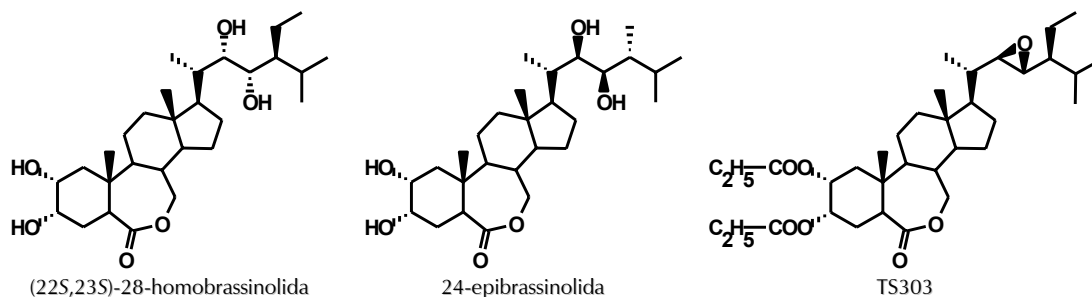


Figura 11: Estructura de la (22S,23S)-28-homobrassinolida, 24-epibrassinolida i TS303.

3.7. Bioassajos.

Des del descobriment dels BRs s'han realitzat nombrosos bioassajos per tal d'avaluar la seva resposta i l'activitat.⁶⁶ Inicialment es van utilitzar bioassajos aplicats en l'estudi d'altres fitohormones.^{46,67} Dels més de 20 bioassajos diferents, els més sensibles han estat el test d'inclinació de la làmina d'arròs (*Rice Lamina Inclination Test* o RLIT)⁶⁸ i el test d'elongació del segon internodi de la mongeta⁶⁹. Aquests han anat evolucionant i adaptant-se a l'estudi dels BRs. Com s'ha comentat, a nivell fisiològic els BRs actius provoquen l'elongació cel·lular i l'augment de la divisió cel·lular, que en el cas del RLIT provoquen una inclinació de la fulla (Figura 12a) i en el cas del test de la mongeta provoquen una elongació del segon internodi (Figura 12b). Ambdós efectes són proporcionals a la resposta dels BRs.

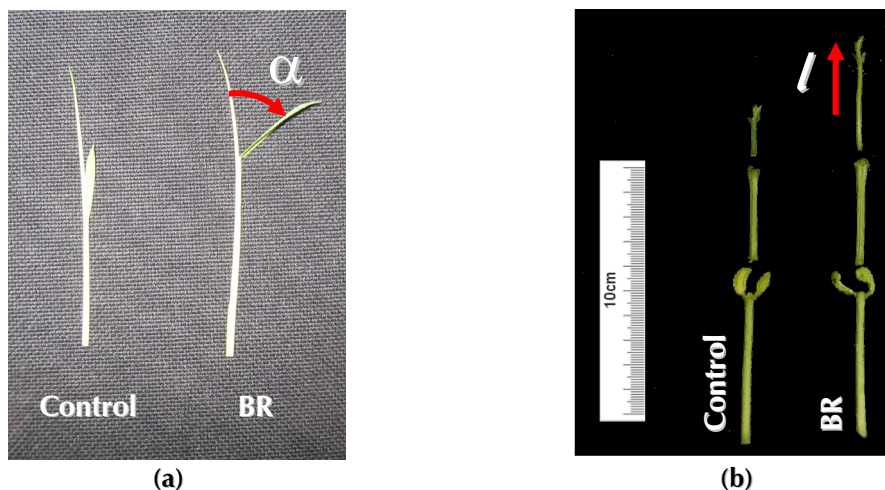


Figura 12: Resposta del test d'inclinació de la làmina d'arròs (a) i del test d'elongació del segon internodi de la mongeta (b).

El test d'inclinació de la làmina d'arròs (RLIT) va ser desenvolupat per Maeda l'any 1965 per estudiar l'efecte de les auxines.⁶⁸ Existeixen dos varietats: a) RLIT amb planta tallada⁷⁰ on els BRs s'apliquen per immersió de la planta tallada en una dissolució de BR, i b) RLIT amb planta sencera on els BRs s'apliquen per injecció a la planta sencera. El test amb planta sencera va ser desenvolupat per Takeno i Pharis l'any 1982.⁷¹

3.8. Estudis qualitius de estructura-activitat (SAR).

S'han realitzat diferents estudis de relació estructura activitat qualitius (SAR). Els resultats són més o menys estrictes en funció del conjunt de BRs analitzats i del bioassaig utilitzat.

La primera aproximació data de l'any 1981.⁷² Wada i Marumo van sintetitzar diversos anàlegs BRs amb la cadena lateral tipus (22S,23S)-28-homoBRs i modificacions a l'anell A. Van suggerir que el diol 2 α ,3 α i la funcionalitat lactona eren essencials per a l'activitat.

Un any més tard (1982) Thompson i col·laboradors⁷³ es van centrar en la influència del substituent a C24 i la configuració del diol en C22,C23. Van establir que els requisits estructurals essencials per a l'activitat eren: a) unió *trans* dels anells A i B, b) presència de lactona o cetona a l'anell B, c) diol 2 α ,3 α , d) diol en C22,C23 amb configuració *cis*, l'èstereoquímica del qual és indiferent i e) un substituent metil o etil en C24 també amb independència de l'èstereoquímica. En l'estudi van utilitzar els dos bioassajos de la mongeta, el del primer internodi i el del segon internodi, i van observar que els perfils de l'activitat biològica eren diferents en cadascun dels bioassajos.

El 1984 Cerana i col·laboradors⁷⁴ van arribar a les mateixes conclusions fent servir els test d'elongació de segments d'arrel de blat.

Entre 1983 i 1984 Takatsuto i col·laboradors⁷⁵⁻⁷⁷ van treballar amb diversos bioassajos. Amb el test d'inclinació de la làmina d'arròs van destacar la importància de que el diol de la cadena lateral fóra 22R,23R i que la resta alquílica en C24 tingués una configuració S. Amb el test del rave i del tomàquet van afirmar que el diol 22R,23R era necessari, que la cetona de l'anell B no era imprescindible i que l'absència de cadena alquílica en C24 era acceptable. No obstant, un cop més, l'activitat relativa observada en cada bioassaig era diferent.

L'any 1992 Yokota i Mori⁷⁸ van afirmar que la diversitat estructural en C24 afectava molt poc als resultats del test d'inclinació de la làmina d'arròs. Cosa que no passava en altres bioassajos.

Un any abans (1991) Kohout i col·laboradors⁶⁶ van sintetitzar alguns compostos sense cadena lateral i amb cadena lateral de pregnà o androstà. També van sintetitzar èsters i amides a partir d'esquelets d'androstà i colestà. Mentre que alguns compostos van resultar ser actius en test del segon internodi de la mongeta, no van donar activitat ni en el test del primer internodi de la mongeta ni en el d'inclinació de la làmina d'arròs.

Malgrat les particularitats de cada estudi, els requisits estructurals acceptats al 1996 eren:

- 2 α ,3 α -diol en l'anell A.
- Funcionalitat lactona o cetona en l'anell B, sent la primera millor.
- Unió dels anells A/B–*trans*.
- Diol *cis* en C22,C23 preferentment amb configuració 22*R*,23*R*.
- Substituent metil o etil en C24.

4. Brassinoesteroides a l'Institut Químic de Sarrià (IQS).

Des de ja fa dos dècades, el laboratori d'esteroides de l'IQS treballa en el camp dels BRs.⁷⁹ La recerca s'ha abordat de forma multidisciplinària realitzant-se estudis de síntesi de BRs tan naturals com no naturals, estudis d'activitat biològica tan a nivell de laboratori com de proves a camp, estudis de correlació estructura activitat quantitativs mitjançant tècniques de modelització molecular, etcètera.

4.1. Síntesi de BRs naturals.

Durant la primera dècada es va dedicar un gran esforç a millorar la síntesi dels BRs naturals 24-epibrassinolida i 28-homobrassinolida prenent com a productes de partida l'ergosterol i estigmasterol respectivament.^{80,81}

El punt clau en la síntesi de BRs és la introducció del diol en C22,C23 (Figura 13). La dihidroxilació del doble enllaç Δ^{22} amb tetròxid d'osmi dona una barreja dels diastereoisòmers 22*R*,23*R* i 22*S*,23*S* sent aquest últim majoritari.⁸² La dihidroxilació asimètrica catalitzada amb tetròxid d'osmi fent servir com a lligand quiral dihidroquinidin 9-O-(9'-fenantril) èter (DHQD-PHN) va augmentar considerablement la proporció del diastereoisòmer 22*R*,23*R*.⁸³ Així en la síntesi de la 28-homobrassinolida sense lligand quiral la proporció entre 22*R*,23*R* i 22*S*,23*S* és de 1 a 24, mentre que amb lligand quiral és de 2,6 a 1.⁸³ En la síntesi de la 24-epibrassinolida la proporció entre 22*R*,23*R* i 22*S*,23*S* és de 1 a 1,8 sense lligand quiral, i de 19,5 a 1 amb lligand quiral.⁸⁴

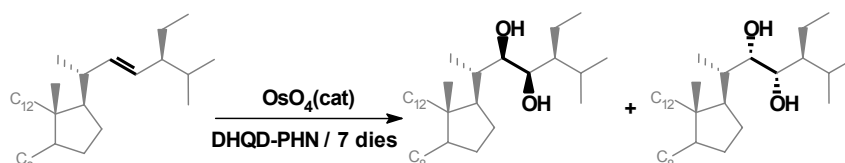


Figura 13: glicolització asimètrica de la cadena lateral tipus *homo*.

4.2. Síntesi de BRs no naturals.

A l'equip s'han sintetitzat nombrosos anàlegs amb modificacions estructurals en diferents parts de la molècula: Anell A, anell B i cadena lateral. A l'anell A les modificacions s'han centrat en el tipus de funcionalitat polar, la seva posició en l'anell i la seva estereoquímica. També s'han sintetitzat BRs amb la unió dels anells A/B *cis*. A l'anell B el nombre de modificacions ha estat més reduït, i s'han centrat en el tipus de funcionalitat polar present a C6. En la cadena lateral, exceptuant la síntesi dels diastereoisòmers 22*S*,23*S*, el nombre de

modificacions és molt reduït i es centra en la síntesi de la cadena lateral saturada o amb un sol hidroxil amb cadenes laterals del tipus 28-homo i 24-epi.

L'estudi de la diversitat estructural de la cadena lateral s'ha optat per fer-lo a través dels que s'anomenen anàlegs BRs androstànics.⁸⁵

4.3. Bioassaig.

La manca de homogeneïtat de les dades d'activitat trobades a la bibliografia van fer que s'iniciés l'estudi de diferents protocols per disposar d'un bioassaig en el propi laboratori. La primera aproximació es va fer amb cultius de cèl·lules d'api, però els resultats no eren reproduïbles.^{86,87} Posteriorment, es va fer servir el test d'inclinació de la làmina d'arròs tan amb planta sencera com amb planta tallada^{84,88}. Amb planta sencera es va seguir inicialment, el procediment descrit per Takeno i Pharis⁷¹ fent servir la varietat japonesa d'arròs nan *Tan-ginbonzu*, però l'equip va aconseguir adaptar el bioassaig a una varietat autòctona del Delta de l'Ebre, la varietat *Bahia*, i millorar-ne la sensibilitat.^{84,89,90} El protocol d'aquest test s'explicarà detalladament en el quart capítol de la discussió de resultats (pàgina 197).

Per determinar el valor d'activitat, es realitzen assajos a diferents dosi per construir la corba que relaciona el logaritme de la dosi amb la resposta. El valor d'activitat és correspon amb la concentració de BR a la qual la resposta és el 50% de la resposta màxima ($-\log(dosi_{50\%})$ o LD_{50}). No obstant, degut a que per alguns BRs no és possible determinar la resposta màxima i tenint en compte que pels BRs actius el 50% de la resposta màxima correspon, en promig, a 45° es pren com a valor d'activitat $-\log(dosi_{45^\circ})$.

En un altre escala, els anys 1994 i 1996 es van fer proves a camp aplicant tan BRs naturals com sintètics en camps de gira-sol,^{84,91} ordi i colza.⁹⁰

4.4. Estudi d'anàlegs BRs androstànics.

Com s'ha dit anteriorment, el pas limitant en la síntesi de BRs és la glicolitació asimètrica de la cadena lateral. És necessari l'ús d'un lligand quirall per obtenir l'èstereoisòmer desitjat. L'elevat preu d'aquest lligand quirall, el qual està en proporcions equimolars, fan que aquest pas de síntesi sigui, econòmicament, molt car. L. Kohout del "Institute of Organic Chemistry and Biochemistry of the Academy of Science" de la República Txeca va proposar com alternativa als BRs naturals 1) anàlegs d'androstà en els quals la cadena s'uneix a l'esquelet mitjançant una esterificació (*Figura 14a*) i 2) anàlegs on l'esquelet s'obté a partir de la degradació de la cadena de l'estigmasterol i la cadena s'uneix mitjançant la formació d'una amida (*Figura 14b*).⁹² A més a més, aquest tipus d'anàlegs obren les portes a l'estudi de l'extrem final alquílic de la cadena lateral.

L'any 1993 Es va iniciar una estreta col·laboració entre el laboratori d'esteroides de l'IQS i l'equip dirigit per L. Kohout, en la qual el laboratori d'esteroides s'havia d'encarregar d'avaluar l'activitat biològica dels compostos sintetitzats a Praga i fer-ne la modelització molecular. Els anàlegs **1** i **2** (*Figura 14*) no van resultar actius en el bioassaig del test d'inclinació de la làmina d'arròs (RLIT) realitzat a l'equip,⁹⁰ però si en el del segon entrenús de la mongeta realitzat a la República Txeca i en proves de camp.^{66,93,94} Es van iniciar estudis computacionals per avaluar la qualitat d'aquests i altres anàlegs no sintetitzats com a potencials BRs.^{95,96} Aquests es mostraran en detall en el primer capítol de la discussió de resultats (pàgina 57) i són el punt de partida sobre el qual es desenvoluparà tot aquest capítol. Actualment, la síntesi dels nous anàlegs BRs androstànics proposats a partir dels estudis de modelització molecular, es realitza conjuntament entre tots dos laboratoris.

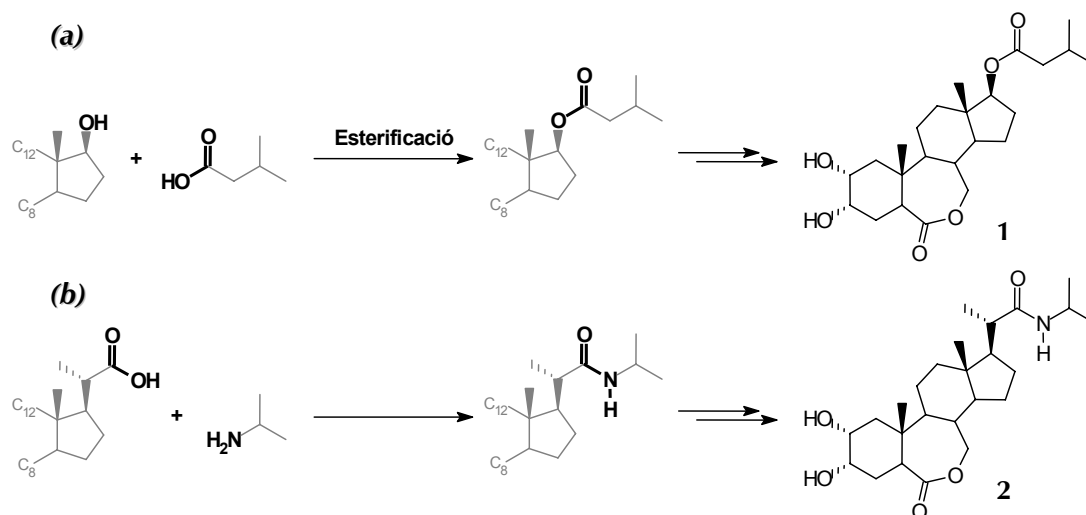


Figura 14: Alternatives a la síntesi de BRs. Construcció de la cadena lateral i exemple d'anàlegs.

4.5. Relacions estructura-activitat quantitatives (QSAR).

Es van iniciar estudis de relació estructura-activitat ja que es considerava que els requisits estructurals postulats al 1996 distaven molt de ser generals degut a que havien estat establerts amb un reduït nombre d'anàlegs BR que presentaven poques modificacions estructurals. Es va observar que: a) al contrari del que proposava la bibliografia, les funcionalitats no podien considerar-se independents, ja que aquestes estaven estretament relacionats entre si, i b) els requisits postulats limitaven el marge d'aplicació a unes poques funcionalitats independents.⁹⁷ Per corroborar la feblesa dels requisits postulats, l'equip va sintetitzar dos anàlegs amb dos requisits contraris als descrits (diol 2β3β i unió A/B cis). Tots dos anàlegs van resultar ser actius.^{98,99} Per tant, es va considerar necessari realitzar nous estudis de correlació estructura-activitat, fent servir tècniques de modelització molecular per tal de tenir en compte l'estructura tridimensional de les molècules i la seva distribució de cargues total.

D'acord amb els estudis de Li¹⁰⁰ i Chory¹⁰¹, es va partir de la base que els BRs actuen a nivell molecular mitjançant la formació d'un complex substrat-receptor que regula l'expressió d'uns gens específics.¹⁰² En el moment d'iniciar els estudis, es desconeixia el receptor biològic dels BRs i en conseqüència es van aplicar mètodes indirectes. La seva utilització porta implícita la necessitat d'haver de considerar que en l'estructura de les molècules d'estudi hi ha inclosa tota la informació necessària per explicar el seu comportament, que totes les molècules s'uneixen a un únic receptor i que ho fan mitjançant els mateixos punts d'unió.

La unió entre el substrat i el seu receptor es dona en una conformació preferent que s'anomenarà "*conformació activa*". En aquesta conformació, els àtoms implicats directament en la unió han de tenir la mateixa distribució espacial en totes les molècules actives. Com la brassinolida és el BR natural més actiu s'assumeix que la seva conformació activa és la que presenta una distribució espacial òptima per unir-se al receptor. Per tant, es pren la conformació activa de la brassinolida com la conformació de referència sobre la qual es comparen les conformacions actives de la resta de BRs i anàlegs. Quan més similar sigui la conformació de l'anàleg d'interès a la conformació activa de la brassinolida, més actiu serà. Aquesta idea es coneix amb el nom d'aproximació a l'anàleg actiu (AAA),¹⁰³⁻¹⁰⁵ i és la base dels estudis SAR realitzats a l'equip.^{95,97}

Alhora de definir la conformació activa, els BRs presenten el problema afegit que són molècules flexibles i això obliga a realitzar una anàlisi conformacional prèvia.

Per desenvolupar els estudis estructura-activitat, l'equip va partir d'un ampli conjunt de BRs amb suficients modificacions estructurals, l'activitat dels quals havia estat determinada pel propi equip mitjançant el test d'inclinació de la làmina d'arròs. Mitjançant tècniques de modelització molecular es va realitzar una anàlisi conformacional, seguida de la selecció de la conformació activa sobre la qual es van calcular tot un conjunt de descriptors moleculars. Aquests es van correlacionar amb l'activitat de forma qualitativa (SAR) i quantitativa (QSAR) fent servir diverses tècniques estadístiques.^{95,97} El desenvolupament d'aquests estudis es mostra esquemàticament a la *Figura 15*. Pas a pas, les diferents etapes són:

- 1. Anàlisi conformacional:** es va realitzar fent servir un anàlisi sistemàtic en forma d'arbre, el qual s'explicarà detalladament en el segon capítol de discussió de resultats (pàgina 93), i es va corroborar mitjançant dinàmica molecular.^{95,97,106,107}
- 2. Selecció de la conformació activa:** es basa en l'aproximació a l'anàleg actiu prenent com a referència la brassinolida. La comparació estructural es va realitzar fent servir dos criteris geomètrics (Sum of squares (SQR) i root mean squares (RMS)) i un criteri volumètric (molecular volume overlay (MVO)).^{95,97,106,107} Aquests criteris, així com tot el procés de selecció de la conformació activa dels BRs s'explicaran detalladament en el segon capítol de discussió de resultats (pàgina 94).

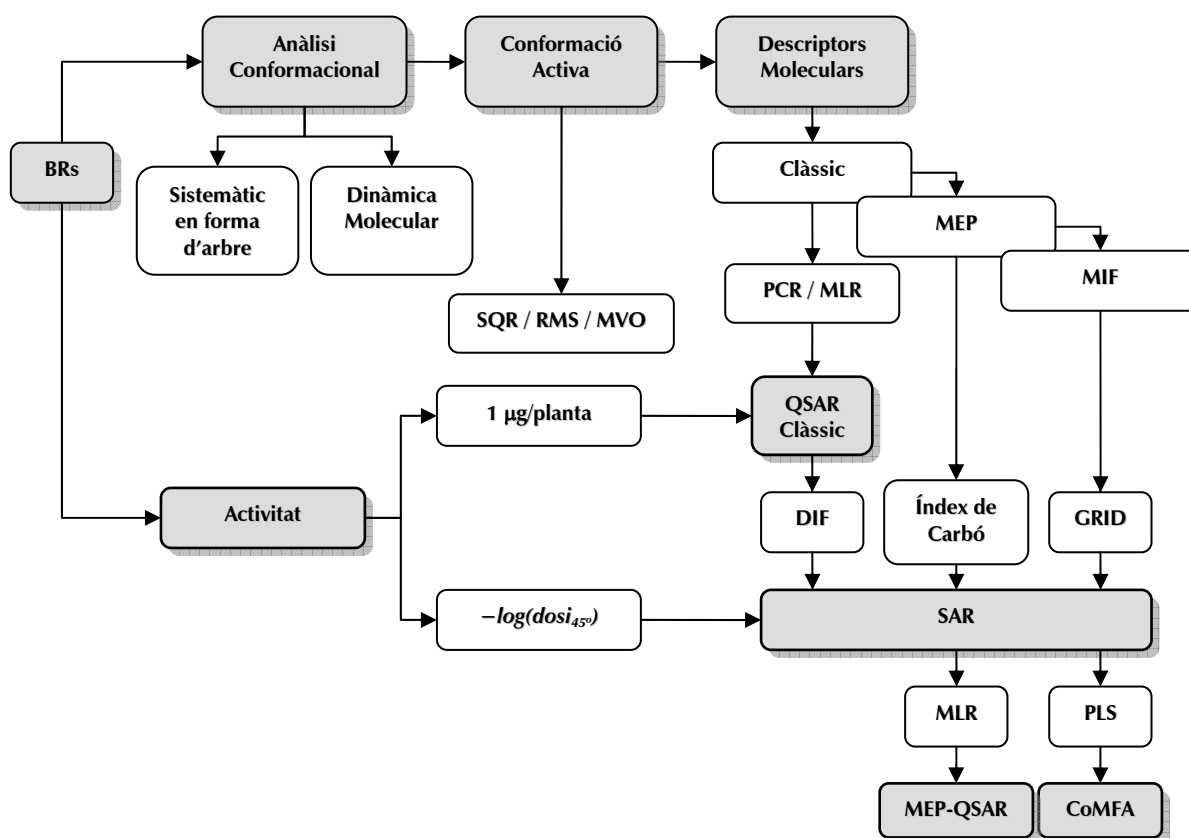


Figura 15: Desenvolupament de la metodologia utilitzada en l'equip per als estudis estructura-activitat en BRs.

- 3. Primers estudis de QSAR (QSAR clàssic):** van considerar la resposta a 1 µg/planta enlloc de l'activitat pròpiament dita. Primer, es va realitzar un anàlisi de Free-Wilson¹⁰⁸⁻¹¹¹ per tenir una idea preliminar de quines funcionalitats eren les més importants. El resultat va ser que el diol 2α,3α de l'anell A participava en un 25%, el

carbonil de l'anell B amb un 11% i el diol 22*R*,23*R* de la cadena lateral amb un 35%.^{99,112} Posteriorment, es van utilitzar descriptors hidrofòbics ($\log P$, ΔH_{solv}), electrostàtics (moment dipolar) i estèrics (volum molecular, moments d'inèrcia). Aquests es van intentar correlacionar amb l'activitat mitjançant regressió lineal múltiple (MLR), regressió de component principals (PCR) i l'ús d'autocorrelogrames. Es va observar com l'activitat correlacionava amb les distàncies entre els àtoms d'oxigen. L'equip va definir un índex de similitud geomètric (anomenat DIF) com la suma de la diferència absoluta entre els valors de cadascuna de les distàncies entre els àtoms d'oxigen del BR en relació amb els de la brassinolida. Quan més baix és aquest índex per un determinat compost, més similar és a la brassinolida i en conseqüència més actiu.^{99,106,112}

4. **Segon estudi de QSAR (MEP-QSAR):** sobre la base que els oxígens tenen un paper molt important, es va triar com a descriptor el potencial molecular electrostàtic (MEP) descrit a través de l'índex de similitud electrostàtic de Carbó (CI),¹¹³ prenent un cop més la brassinolida com a referència. Quan major és aquest índex, més similar és un compost a la brassinolida i en conseqüència és més actiu. A partir d'aquest moment es va començar a utilitzar el valor adequat d'activitat, expressat a través del $-\log(\text{dosi}_{45\%})$. Qualitativament es va observar una bona correlació entre la superposició del potencial molecular electrostàtic i l'activitat.^{95,97} Quantitativament i mitjançant la regressió lineal múltiple (MLR) es va obtenir una bona correlació entre l'índex de similitud electrostàtic de Carbó (CI) i l'activitat.⁹⁵
5. **Tercer estudi de QSAR (CoMFA):** veient que la utilització de descriptors tridimensionals (descriptors que consideren la propietat estructural no com un valor, sinó com una distribució en l'espai) rendia bons resultats, es va optar per fer servir com a descriptors els camps d'interacció molecular (MIF). Quantitativament, primer es van reduir els MIF a uns índex de similitud, però degut a la reducció d'una propietat 3D a uns índex les correlacions obtingudes van ser difícils d'interpretar en termes de funcionalitats químiques. Després, es va realitzar un anàlisi comparatiu de camps moleculars (CoMFA) mitjançant mínims quadrats parcials (PLS). Les correlacions presentaven el risc d'obtenció de correlacions a l'atzar. Finalment, els millors resultats es van obtenir amb la metodologia GRID-GOLPE, també anomenada CoMFA-like. Amb aquesta metodologia es va obtenir un model amb bons paràmetres estadístics i una bona interpretació en termes de funcionalitats químiques.⁹⁵

La utilització de camps molecular va permetre estudiar el tipus d'interacció que podia tenir lloc en la unió BR-receptor. Es va observar que les interaccions per enllaç d'hidrogen jugaven un paper molt important tan quantitativa com qualitativament.⁹⁵⁻

⁹⁷ Qualitativament i prenent la brassinolida com a referència, es va observar una bona correlació entre la superposició dels mapes de GRID obtinguts amb la sonda d'aigua i l'activitat. A la *Figura 16* es mostra el mapa de GRID* de la brassinolida a -3 kcal/mol fent servir la sonda d'aigua. Aquest mostra les zones on es poden formar enllaços per enllaç d'hidrogen.

* El càlcul dels camps de GRID s'explica detalladament en el capítol de teoria i mètodes (pàgina 47).

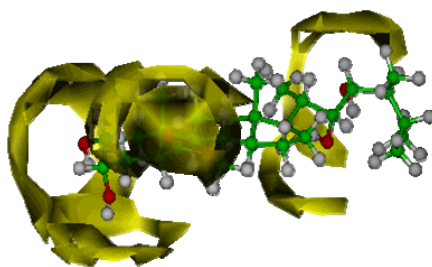


Figura 16: Mapa de GRID de la brassinolida a -3 kcal/mol fent servir la sonda d'aigua.

- 6. Estudis qualitius estructura-activitat (SAR):** com s'ha dit anteriorment, per cadascun dels descriptors (DIF, MEP i GRID) es van establir un conjunt de relacions estructura-activitat qualitatives (SAR) en cadascuna de les tres zones d'influència dels BRs: l'anell A, l'anell B i la cadena lateral.^{95,97} Els BRs considerats per estudiar cadascuna de les zones contenen exclusivament modificacions estructurals en aquella zona d'influència, es a dir, les altres dos zones són comunes per a tots els compostos.

a) Anell A.

A la *Figura 17* es mostren: a) les modificacions estructurals presents en l'anell A juntament amb l'estructura comuna de l'anell B i la cadena lateral, b) l'estructura del compost de referència, la brassinolida, c) el valor del paràmetre DIF, d) la superposició dels potencials moleculars electrostàtics a -30 KT entre el BRs (en vermell) i la brassinolida (en groc), e) el valor de l'índex de similitud electrostàtic de Carbó, f) la intersecció dels mapes de GRID, es a dir, la zona comuna en la superposició dels mapes de GRID* amb la sonda d'aigua, i g) el valor d'activitat expressat com a $-\log(dosi_{45\%})$.

En la sèrie d'anàlegs de l'anell A no hi ha correlació entre el paràmetre geomètric DIF i l'activitat. Tampoc hi ha correlació entre l'índex electrostàtic de Carbó (CI) i l'activitat. No obstant, qualitativament s'observa que a mesura que l'àrea corresponent al MEP de l'hidroxil en 3 (zona en vermell) s'apropa a la de l'hidroxil en 3α de la brassinolida (zona en groc) l'activitat augmenta. Quelcom similar passa amb la intersecció dels mapes de GRID, a mesura que aquesta augmenta, l'activitat també augmenta.

La manca de correlació dels paràmetres DIF i CI es va atribuir al fet que petits canvis en una part de la molècula afecten poc a tot el conjunt, i tan DIF com CI han estat calculats tenint en compte tota la molècula i no només l'àrea d'estudi.

* La zona d'intersecció entre dos mapes de GRID es va estimar a través de la suma dels dos camps i l'aplicació d'un tall d'energia de -6 kcal/mol.

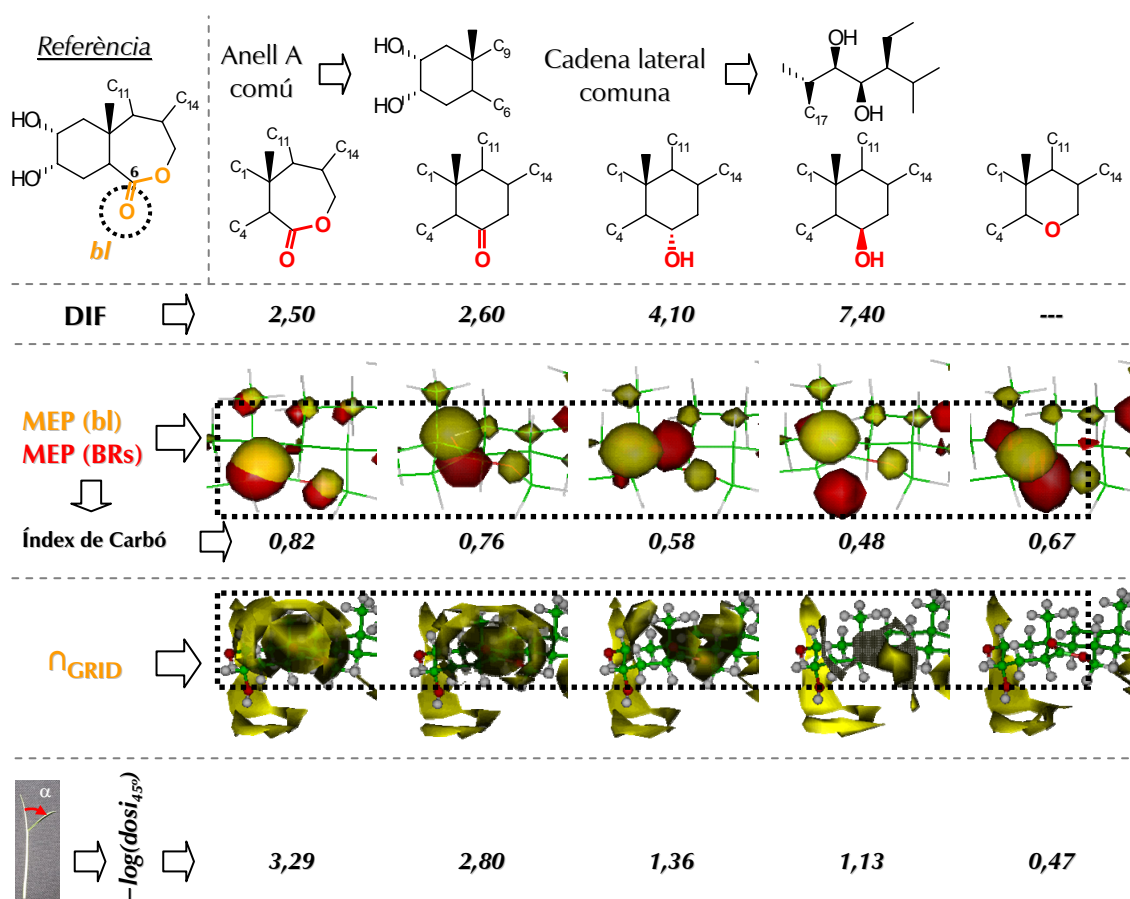


Figura 17: Resultats del SAR en l'anell A per els descriptors DIF, superposició del MEP i intersecció dels mapes de GRID. Es remarca la zona corresponent l'hidroxil en 3α de la bl.

b) Anell B.

De forma similar al cas anterior, a la *Figura 18* es mostren: a) les modificacions estructurals presents en l'anell B juntament amb l'estructura comuna de l'anell A i la cadena lateral, b) l'estructura del compost de referència, la brassinolida, c) el valor del paràmetre DIF, d) la superposició dels potencials moleculars electrostàtics a -30 KT entre el BRs (en vermell) i la brassinolida (en groc), e) el valor de l'índex de similitud electrostàtic de Carbó, f) la intersecció dels mapes de GRID, es a dir, la zona comuna en la superposició dels mapes de GRID amb la sonda d'aigua, i g) el valor d'activitat expressat com a $-\log(dosi_{45^\circ})$.

En la sèrie d'anàlegs de l'anell B s'observa una correlació entre el paràmetre geomètric DIF i l'activitat: A mesura que DIF disminueix, més actiu és un BRs ja que la relació de distàncies entre els àtoms són cada cop més similars a les de la brassinolida. Amb l'excepció de l'èter, també s'observa una correlació entre l'índex electrostàtic de Carbó (CI) i l'activitat: A mesura que CI augmenta, més actiu és un BRs ja que l'àrea corresponent al MEP de la funció oxigenada en C6 (zona en vermell) s'apropa a la del carbonil en C6 de la brassinolida (zona en groc). Quelcom similar passa amb la intersecció dels mapes de GRID, a mesura que aquesta augmenta, l'activitat també augmenta. A més a més, tot sembla indicar que la part alta del carbonil en C6 juga el paper més important en aquesta correlació.

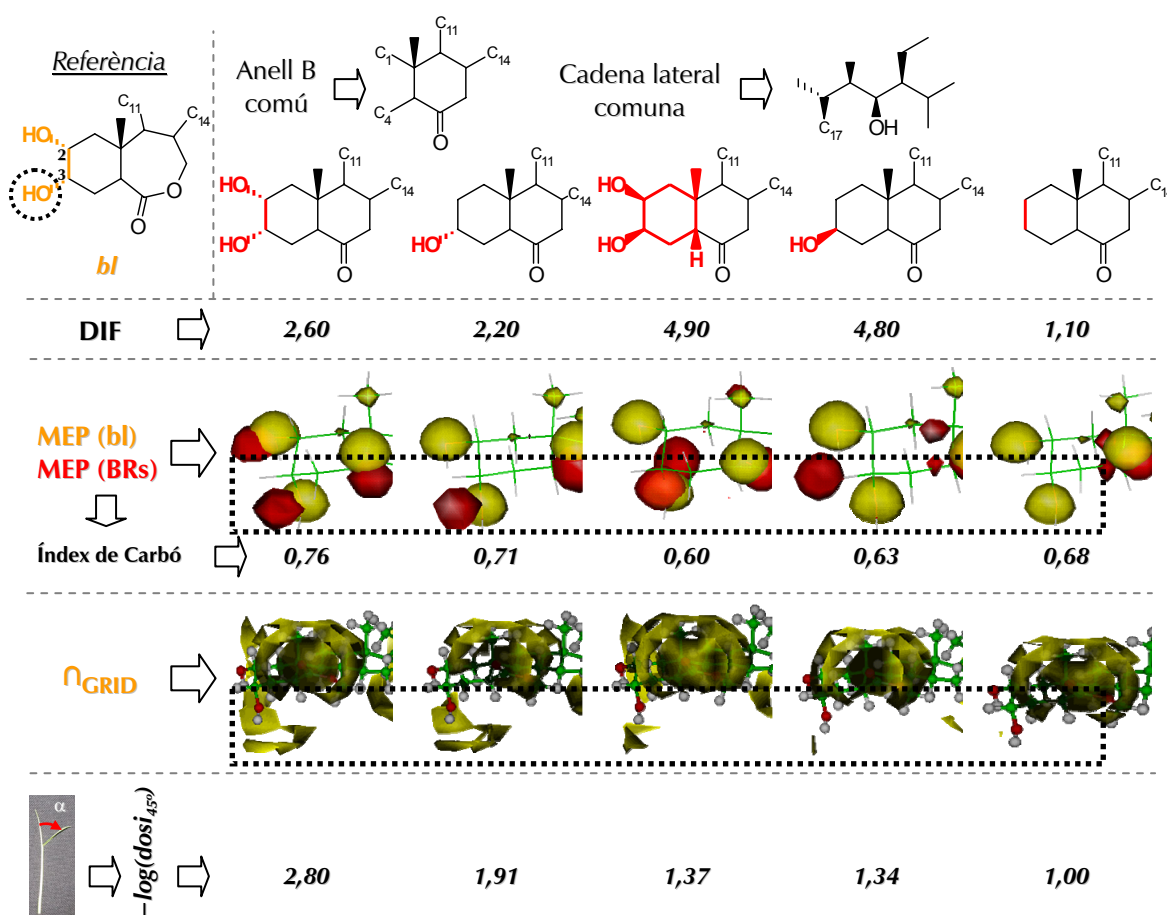


Figura 18: Resultats del SAR en l'anell B per els descriptors DIF, superposició del MEP i intersecció dels mapes de GRID. Es remarca la zona corresponent al carbonil en C6 de la bl.

c) Cadena lateral.

A la Figura 19 es mostren: a) les modificacions estructurals presents en la cadena lateral juntament amb l'estructura comuna dels anells A i B, b) l'estructura del compost de referència, la brassinolida, c) el valor del paràmetre DIF, d) la diferència a -30 KT entre els potencials moleculars electrostàtics de la brassinolida i els compostos, e) el valor de l'índex de similitud electrostàtic de Carbó, f) la superposició de mapes de GRID entre els BRs (en groc) i la brassinolida (en vermell), i g) el valor d'activitat expressat com a $-\log(dosi_{45^\circ})$.

Igual que passava en l'anell A, en la sèrie d'anàlegs de la cadena lateral, tampoc hi ha correlació amb l'activitat ni del paràmetre geomètric DIF ni de l'índex electrostàtic de Carbó (CI). Novament la falta de correlació es va argumentar pel fet que tan DIF com CI són descriptors moleculars globals.

Pel que respecta a les superposicions dels mapes de GRID, s'observa com a mesura que augmenta la superposició en la zona corresponent a la part superior de l'hidroxil en C23 de la bl, major és l'activitat dels compostos. Per a la 22S,23S-homobrassinolida (S_hbl) no hi ha superposició entre els dos hidroxils, però la zona d'interacció del 23S-OH no queda gaire lluny de la del 23R-OH de la brassinolida. La seva activitat és 1,78. Per la 22S,23S-epibrassinolida (S_epibl) s'observa una petita superposició en l'extrem més superior de l'hidroxil en C23. En conseqüència la seva activitat augmenta a 2,08 (+0,30 respecte l'anterior). Per a la 24-epibrassinolida (epibl) la superposició en C23 és gairebé total, però es troba lleugerament desplaçada

cap a la dreta. La seva activitat augmenta a 2,45 (+0,37 respecte l'anterior). Per a la 28-homobrassinolida (hbl) la superposició en C23 és gairebé total. La seva activitat augmenta considerablement a 3,29 (+0,84 respecte l'anterior).

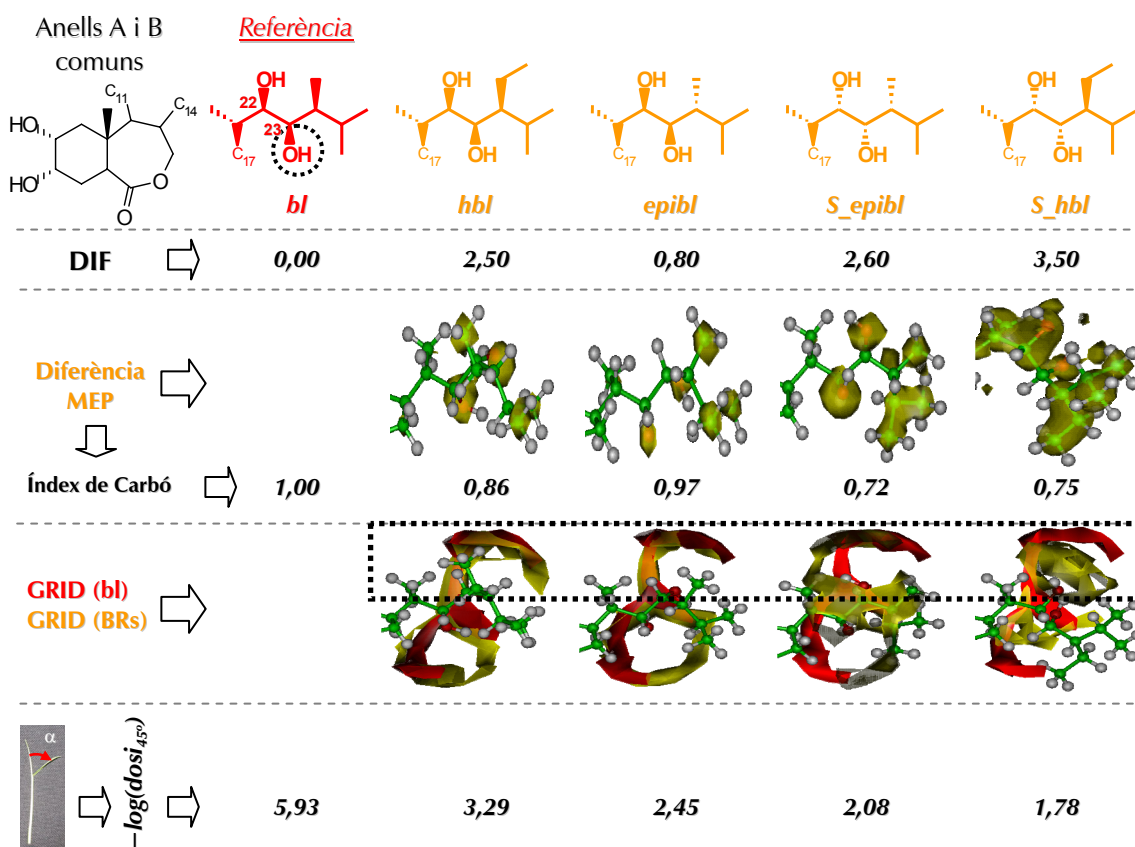


Figura 19: Resultats del SAR en la cadena lateral per els descriptors DIF, diferència del MEP i superposició dels mapes de GRID. Es remarca la superposició en la part superior corresponent a l'hidroxil en C23 de la bl.

Vist que la formació d'enllaços d'hidrogen per part dels hidroxils sembla tenir un paper fonamental en l'activitat dels BRs, paral·lelament a aquesta tesi, s'han realitzat diversos estudis per esbrinar el mode d'actuació d'aquests hidroxils. Recentment, s'han canviat alguns dels hidroxils de la 28-homocastasterona i la 28-homobrassinolida per grups acceptors d'enllaç d'hidrogen com són fluor,¹¹⁴⁻¹¹⁸ cetona¹¹⁷ i azida,^{115,117,119,120} i per grups dadors d'enllaç d'hidrogen com és l'amina.¹¹⁷ A partir de les dades d'activitat, sembla ser que els hidroxils del diol 2 α ,3 α de l'anell A actuen com acceptors d'enllaç d'hidrogen i es suggereix que els hidroxils del diol 22R,23R de la cadena lateral poden actuar com a donadors d'enllaç d'hidrogen.^{117,121}

- Estudi computacional del KM-01:** l'any 1994 Kim i col·laboradors van aïllar dels fongs *Dreschlera avenae* i *Pycnoporus coccineus* un compost eremofilànic (KM-01, Figura 20) que inhibeix específicament l'acció dels BRs.¹²² Sota la hipòtesi que el KM-01 actua com antagonista de la brassinolida i que en el receptor competeix per els mateixos llocs d'unió, l'equip va iniciar el seu estudi computacional.^{95,97} Basant-se en l'aproximació a l'anàleg actiu, es va comparar amb la brassinolida i es va realitzar un estudi de la selectivitat de les interaccions fent servir la metodologia descrita per Cruciani i Goodford.¹²³

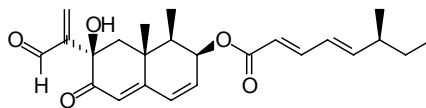


Figura 20: Estructura del KM-01.

Tot i tenir una estructura molt diferent a la brassinolida, el carbonil de l'aldehid del KM-01 es superposa amb l'hidroxil en 2 α de la brassinolida, igualment la cetona es superposa amb el carbonil a C6, el carbonil de l'èster s'orienta cap a l'hidroxil en C22 i les cadenes laterals estan molt properes entre si.

Aquest estudi juntament amb els resultats del SAR, han donat lloc a la idea que els hidroxils 2 α i 22R són els responsables de la unió, mentre que els hidroxils 3 α , 23R i la part alta del carbonil a C6 són els responsables de desencadenar l'activitat.⁹⁷

- 8. Estudis quantitius estructura-activitat (QSAR):** la interpretació dels models CoMFA i GRID-GOLPE va donar peu a identificar no només les zones que contribuïen positiva i negativament a l'activitat sinó també el tipus d'interacció present. Aquestes es resumeixen esquemàticament en la Figura 21.⁹⁵ S'observa com en els anells A i B els efectes coincideixen força bé amb les conclusions del SAR. En la cadena lateral també, però es considera que la influència tan del reste alquílic en C24 com de l'isopropil terminal és molt genèrica i caldria estudiar-la amb més profunditat. De fet, aquest és l'objectiu principal d'aquesta tesi.

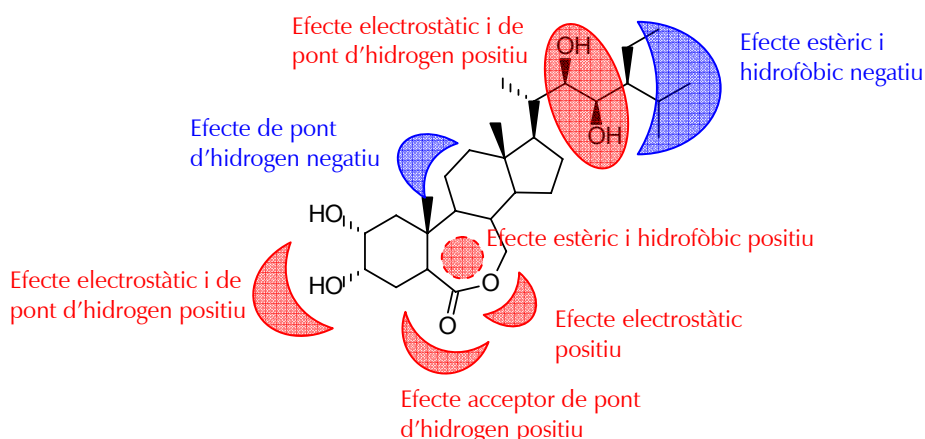


Figura 21: Esquema de les contribucions positives i negatives a l'activitat dels estudis QSAR mitjançant CoMFA i GRID-GOLPE.

Malgrat la bona correlació establerta, cap dels models realitzats fins al moment era capaç d'explicar l'activitat dels dos BRs més actius, la brassinolida i la castasterona.

- 9. Conformació activa basada en mapes de GRID flexibles:** donada la bona correlació entre l'activitat i la possibilitat de formar enllaços per enllaç d'hidrogen, es va proposar fer servir els mapes de GRID considerant la cadena lateral flexible per definir una nova conformació activa.⁹⁵ En la Figura 22 s'observa com la zona d'interacció del diol en C22,C23 dels BRs augmenta al considerar la cadena lateral flexible.

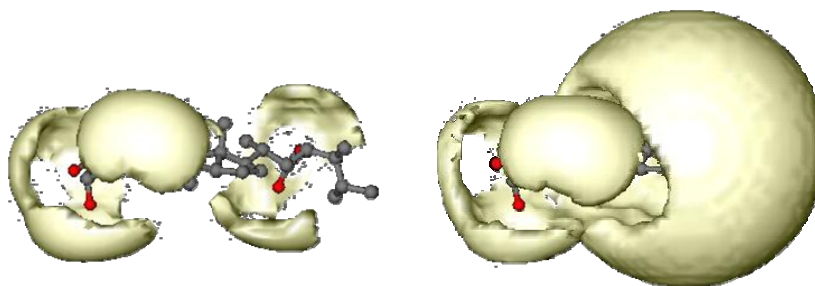


Figura 22: Mapes de GRID rígid i flexible de la brassinolida a -3 kcal/mol.

L'energia en un punt no només dóna idea de la força de la interacció, sinó també de la capacitat de moviment de la cadena. Per tant, fent servir aquests camps, no és necessari definir una conformació activa, ja que el propi camp té en compte el moviment conformational. Aquesta nova metodologia comportava un gran avantatge respecte l'anterior. La metodologia amb tots els seus avantatges, el protocol, el desenvolupament pas a pas i la discussió dels resultats seran explicats detalladament en el segon capítol de la discussió de resultats (pàgina 99). Els models QSAR obtinguts amb aquesta conformació activa van ser capaços d'explicar l'activitat de la brassinolida i la castasterona.

Les diferències entre les dos conformacions actives es van considerar mínimes (*"resultados convergentes"*).⁹⁵

Aquest és el punt de partida d'aquesta tesi, la qual es planteja inicialment com una continuació dels estudis iniciats en l'equip en el camp de les relacions estructura-activitat dels BRs tan a nivell qualitatiu (SAR) com quantitatiu (QSAR) mitjançant mètodes computacionals.

A l'equip es creu que les modificacions estructurals tan en l'anell A com en l'anell B són suficients com per caracteritzar correctament aquestes regions. En canvi, en la cadena lateral les modificacions estructurals són majoritàriament degudes a la combinació de dos factors, la configuració del diol (22*R*,23*R* o 22*S*,23*S*) i el reste alquílic en C24 (*S*-metil, *S*-etil o *R*-metil). Però, mentre que a nivell qualitatiu la influència del diol està ben caracteritzada, no passa el mateix amb el reste alquílic en C24, el qual és responsable dels canvis més importants en l'activitat. Per aquest motiu, l'objectiu principal és centrar-se en l'estudi computacional de la cadena lateral. No obstant, també s'estarà obert a tot allò que amplii els estudis anteriors i/o que aportï una novetat al camp de les relacions estructura-activitat dels BRs.

L'estudi computacional dels anàlegs BRs androstànics incideix plenament en l'objectiu principal d'estudiar el comportament de la cadena lateral i representa una novetat en les relacions estructura-activitat dels BRs. Aquest és el punt de partida d'aquesta tesi i es desenvoluparà en el *"Capítol 1"*.

L'estudi computacional de la cadena lateral començarà volent determinar la influència del reste alquílic en C24 i de cadascun dels hidroxils per separat de la cadena lateral, i acabarà redefinint la conformació activa dels BRs en el *"Capítol 2"* i estudiant el seu impacte sobre els models de QSAR en el *"Capítol 3"*.

La resta d'estudis romanen inicialment oberts. Fent ús d'aquesta llibertat, en el *"Capítol 4"* s'ha determinat l'error experimental de les dades d'activitat.

5. Nomenclatura dels BRs.

Pel que fa a la nomenclatura dels BRs, la nomenclatura IUPAC per esteroides^{124,125} rendeix noms molt complexos i s'ha adoptat una nomenclatura semisistemàtica. No obstant, en aquesta nomenclatura no hi ha un acord universal.¹²⁶⁻¹²⁸ L'any 2003, es va fer un intent per establir una nomenclatura comuna per els BRs en la XX Conferència en Isoprenoides que es va realitzar a Liberec (República Txeca).¹²⁹ Com a resultat en va sortir una proposta que es va publicar a "Plant Growth Regulation".¹³⁰ No obstant, dos anys més tard en la XXI Conferència en Isoprenoides que es va realitzar a Bialowieza (Polònia) es va poder comprovar com l'intent no havia arrelat i cada grup continuava fent servir la seva pròpia nomenclatura.

Davant aquesta falta d'unanimitat, en aquesta tesi s'ha seguit una nomenclatura pròpia. S'ha optat per una nomenclatura semisistemàtica on s'ha procurat que el nom contingui tota la informació referent a l'estructura dels BRs. El nom complet d'un BR es compona unint els noms de cadascun dels fragments estructurals (Figura 23 i full adjunt) en l'ordre: "Cadena lateral" + "Anell B" + "_" + "Anell A".

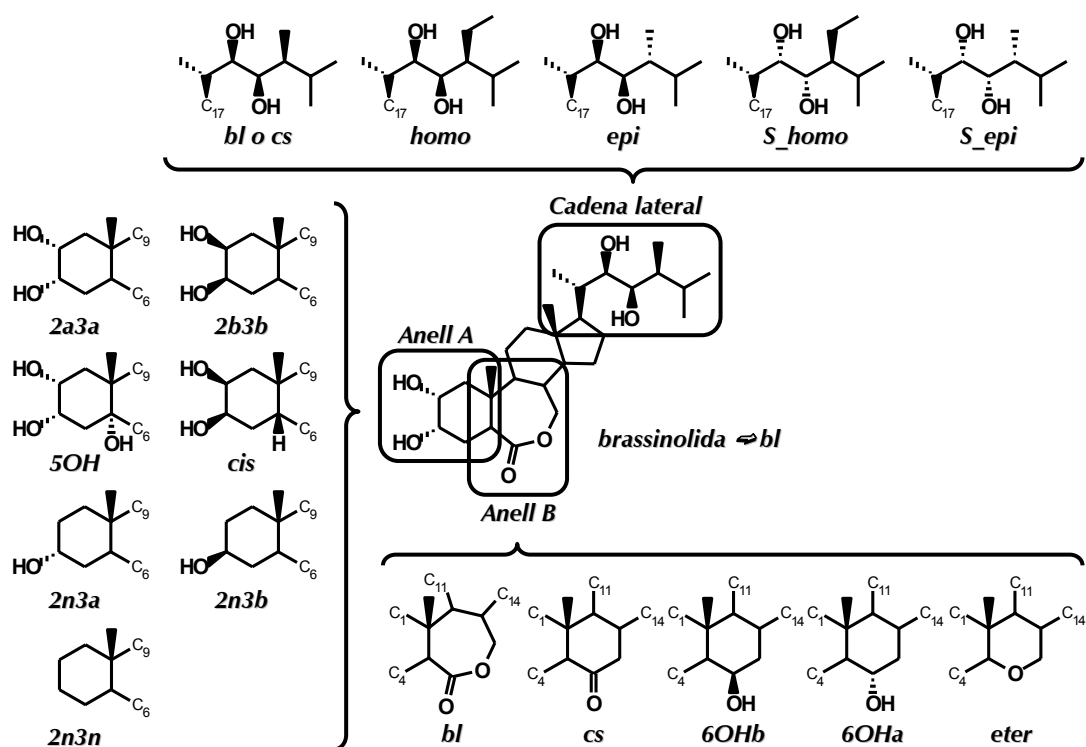


Figura 23: Diversitat estructural dels BRs presents en el QSAR de l'equip i nomenclatura dels fragments estructurals utilitzada en aquesta tesi.

Els fragments estructurals són:

- Cadena lateral:
 - bl o cs: Tan la brassinolida (bl) com la castasterona (cs) es prenen com a estructures de referència. Quan l'estructura és l'original dels BRs (diol 22*R*,23*R* i cadena lateral de campestà) la cadena lateral no té cap nom propi i es passa a anomenar segons l'anell B. Exemple: brassinolida ⇌ bl.
 - homo: És la cadena lateral tipus dels 28-homobrassinosteroides (diol 22*R*,23*R* i cadena lateral d'estigmastà). S'abreia únicament amb la "h". Exemple: 28-homocastasterona ⇌ hcs.

- epi: És la cadena lateral tipus dels 24-epibrassinosteroides (diol 22*R*,23*R* i cadena lateral d'ergostà).
- S_homo: És la cadena lateral tipus dels (22*S*,23*S*)-28-homobrassinosteroides. Igual que els homoBRs s'abreia com "S_h". Exemple: (22*S*,23*S*)-28-homobrassinolida \Rightarrow S_hbl.
- S_epi: És la cadena lateral tipus dels (22*S*,23*S*)-24-epibrassinosteroides.
- Anell B:
 - bl: Es correspon amb la funcionalitat genèrica lactona pròpia de la brassinolida (bl), la qual dóna lloc a un anell B de 7 baules.
 - cs: Es correspon amb la funcionalitat genèrica cetona pròpia de la castasterona (cs), la qual dóna lloc a un anell B de 6 baules.
 - 6OHa: La funcionalitat cetona en C6 ha estat substituïda per un hidroxil amb configuració α . Per indicar que l'anell és de 6 baules s'afegeix a posteriori de "cs" separat per un "_". Exemple: hcs_6OHa.
 - 6OHb: ídem, però amb configuració β .
 - eter: La funcionalitat cetona en C6 ha estat substituïda per una funcionalitat éter. Per indicar que l'anell és de 6 baules s'afegeix a posteriori de "cs" separat per un "_". Exemple: S_hcs_eter.
- Cadena lateral:
 - 2a,3a: El diol 2-alfa,3-alfa (2 α ,3 α) és el més abundant. Per aquest motiu s'omet en el nom dels BRs. En tots els exemples anteriors l'anell A presenta el diol 2 α ,3 α .
 - 2b3b: El diol presenta la configuració 2-beta,3-beta (2 β ,3 β).
 - 2n,3n: El grup funcional no (n) hi és.
 - 2n3a, 2n3b: són combinacions dels casos anteriors. L'hidroxil en C2 no està present mentre que en C3 si que hi és i presenta una configuració alfa o beta.
 - 5OH: Hi ha un hidroxil en C5. Com la unió és A/B trans l'única configuració possible és alfa i per això no s'indica. Seguint la regla anterior, el diol 2 α ,3 α no s'anomena. Exemple: hcs_5OH.
 - cis: La unió dels anells és A/B cis. Com tots els BRs amb unió A/B cis tenen el diol 2 β ,3 β , es sobreentén que tot és un bloc i per això no s'indica la seva configuració. Exemple: hbl_cis.

En el cas dels BRs que no estiguin contemplats, la idea es definir un fragment estructural característic seguint les mateixes idees abans exposades. Per exemple, la cadena lateral saturada s'expressaria com 22n23n donant lloc a la hcs_22n23n o la epics_22n23n (pàgina 187).

OBJECTIUS

Aquesta tesi és una continuació dels estudis iniciats en l'equip en el camp de les relacions estructura-activitat dels brassinoesteroides tan a nivell qualitatiu (SAR) com quantitatiu (QSAR) mitjançant mètodes computacionals.

L'objectiu general d'aquesta tesi és centrar l'atenció d'aquests estudis en la cadena lateral dels BRs. No obstant, s'està obert a tot allò que amplii els estudis anteriors i/o que pugui aportar una novetat al camp de les relacions estructura-activitat dels BRs.

El desenvolupament de l'objectiu principal ha portat a:

- Estudiar les cadenes laterals d'anàlegs BRs androstànics.
- Revisar i redefinir la conformació activa dels BRs posant un especial èmfasi en la cadena lateral.
- Estudiar la influència de la conformació activa dels BRs sobre els models de QSAR.

Fora de l'objectiu principal, però íntimament relacionat amb els estudis de QSAR s'ha volgut:

- Determinar l'error experimental de les dades d'activitat.

TEORIA I MÈTODES

1. Química computacional i/o modelització molecular.

La present tesi s'emmarca dins de l'àrea de coneixement de la *química computacional*. Aquesta és una disciplina de la química que utilitza mètodes matemàtics per a calcular les propietats d'una molècula i/o simular el seu comportament. Inclou metodologies tan diverses com el disseny de síntesi assistit per ordinador, la recerca en bases de dades, la manipulació de llibreries combinatòries, estudis de relació estructura-activitat, etcètera.¹³¹

Sovint, s'utilitza incorrectament el terme de *modelització molecular* com a sinònim de *química computacional*. Un clar exemple són les pàgines web de software en la matèria on aquests dos termes es troben junts, barrejats i confonent-se entre si. La *modelització molecular* és centra en la investigació de l'estructura i/o les propietats moleculars fent servir tècniques de química computacional i visualització molecular per a proporcionar una representació tridimensional creïble donat un conjunt de circumstàncies.¹³¹

De les diverses àrees de la química computacional, aquesta tesi es centra en els estudis de les *relacions estructura-activitat* mitjançant tècniques de modelització molecular.

2. Relacions estructura-activitat.

Es parla de *relacions estructura-activitat* (SAR) quan es pretén establir de forma qualitativa l'associació existent entre aspectes específics de l'estructura molecular i una determinada acció biològica.¹³² Es parla de *relacions estructura-activitat quantitatives* (QSAR) quan es pretén establir mitjançant tècniques de regressió, models matemàtics que relacionen l'activitat biològica d'un conjunt de compostos amb constants fisicoquímiques i/o variables indicadores que depenen de l'estructura dels compostos.¹³²

Fa més d'un segle, l'any 1893, Charles Richet va publicar una nota on relacionava la toxicitat d'una sèrie de compostos orgànics amb les seves propietats fisico-químiques.¹³³ Des de llavors, es van anar desenvolupant diverses teories.¹³⁴⁻¹³⁶ No obstant, el concepte de correlacionar quantitativament activitats i estructures (QSAR) data del 1964 i s'atribueix a C. Hansch i T. Fujita.¹³⁷ Aquest mateix any S. M. Free i J. W. Wilson van desenvolupar un model basat en la contribució additiva de grups als valors d'activitat.¹⁰⁸ Des d'aquesta data, l'anàlisi de Hansch i aproximacions relacionades^{138,139} han tingut un impacte enorme en el desenvolupament de fàrmacs, especialment en l'optimització del cap de sèrie.

En la dècada dels 80 i amb l'avanç dels ordinadors es va anar desenvolupant el concepte de 3D-QSAR.¹⁴⁰ Aquest implica l'anàlisi quantitatiu de la relació entre l'activitat biològica d'un conjunt de compostos i les seves propietats tridimensionals fent servir mètodes estadístics de correlació.¹³¹ Per arribar al seu ple desenvolupament va ser necessari desenvolupar conceptes com la *similitud molecular* i l'*alineament molecular*, descriptors tridimensionals com els *camp moleculars*, mètodes estadístics de correlació, etcètera.

3. Mecànica molecular.

Dins de la modelització molecular hi ha quatre metodologies per estudiar el comportament i les interaccions de les molècules.

- a) *Mecànica quàntica (ab initio)*: Es basa en la utilització de l'equació de Schrödinger sencera per tractar tots els electrons d'un sistema químic. A la pràctica es necessari realitzar aproximacions per restringir la complexitat de les funcions d'ona electròniques i fer el càlcul possible. Els mètodes són independents de qualsevol experiment que no sigui la determinació de les constants fonamentals.¹⁴¹

- b) *Mètodes semi-empírics de mecànica quàntica*: Fan servir paràmetres derivats de dades experimentals per simplificar els càlculs. Aquesta simplificació pot tenir lloc a diversos nivells: Simplificació del Hamiltonià (Extended Huckel method), aproximació a l'avaluació de certes integrals moleculars (Zero differential overlap approximation), simplificació de les funcions d'ona, etcètera.¹⁴¹
- c) *Mecànica molecular*: És un mètode empíric que pretén donar estimacions de les estructures i les energies de les conformacions de les molècules.¹³² Les molècules són tractades fent servir les lleis de la mecànica clàssica.
- d) *Density Functional Theory (DFT)*: Calcula les propietats moleculars en funció de la densitat electrònica enlloc de fer-ho mitjançant el càlcul de la funció d'ona.

En aquesta tesi s'han utilitzat exclusivament mètodes de *mecànica molecular*. Aquests tracten els àtoms i els enllaços que formen les molècules com un sistema de esferes i molles. El *camp de forces (force-field)*¹⁴¹ és el conjunt d'equacions i paràmetres que s'utilitzen per calcular l'energia d'un sistema a partir de les seves coordenades atòmiques. L'expressió del camp de forces contempla: a) un conjunt d'interaccions a través dels enllaços com són les deformacions de la longitud d'enllaç (*stretching*), l'angle d'enllaç (*bending*), el gir o torsió de l'enllaç, les vibracions fora del pla... b) un conjunt d'interaccions a través de l'espai com són les interaccions de Van der Waals i les interaccions electrostàtiques, i c) un conjunt d'interaccions creuades (*cross-terms*) les quals tenen en compte l'acoblament entre els diferents tipus d'interaccions. El camp de forces es complementa amb un conjunt de paràmetres que depenen del tipus d'àtoms (*atom-type*) als quals fa referència. Aquests són etiquetes que s'assignen als àtoms i que pretenen resumir les seves característiques tan estructurals (distància i angles d'enllaç...) com químiques (hibridació, aromaticitat...) tenint en compte l'element químic del que es tracta. La parametrització és l'etapa clau en el desenvolupament d'un camp de forces.

Els camps de forces utilitzats en aquesta tesi són:

- *Consistent Valence Force Field (CVFF)*:^{142,143} Camp de forces general en l'àmbit de la bioquímica. Parametritzat a partir de dades experimentals per tal de reproduir el comportament d'aminoàcids, l'aigua i una gran varietat de grups funcionals. El seu marge d'aplicació compren molècules orgàniques senzilles, pèptids, proteïnes i sistemes orgànics complexes.
- *Tripes Force Field (TriposFF)*:¹⁴⁴ Camp de forces general desenvolupat per la casa comercial Tripos per incloure'l en el seu software (SYBYL). El seu marge d'aplicació compren molècules orgàniques senzilles i proteïnes.
- *Amber89 (Assisted Model Building and Energy Refinement)*:¹⁴⁵⁻¹⁴⁸ Camp de forces general parametritzat a partir de dades experimentals per tal de reproduir el comportament de proteïnes i àcids nucleics. No obstant, ha estat àmpliament utilitzat per a l'estudi de polímers i molècules senzilles.
- *MM94FF (Merck Molecular Force Field)*:¹⁴⁹⁻¹⁵⁵ Camp de forces general considerat de segona generació. Parametritzat a partir de càlculs *ab initio*, el seu marge d'aplicació compren l'estudi de molècules orgàniques senzilles, proteïnes i altres sistemes d'importància biològica.

4. Anàlisi conformacional.

La *conformació* d'una molècula és la col·locació espacial dels àtoms la qual es pot convertir en una nova conformació mitjançant la rotació d'un o diversos enllaços.¹³² Cada conformació porta associada un valor d'energia. L'*espai conformacional* és el conjunt format per totes les conformacions possibles. A efectes pràctics només tenen interès les conformacions que estan en un mínim d'energia, ja sigui local o absolut, i les que caracteritzen un estat de transició.

L'*anàlisi conformacional* consisteix en l'exploració de les conformacions energèticament favorables fent servir tècniques de modelització molecular.¹³¹ L'espai conformacional es pot explorar mitjançant recerques sistemàtiques, aleatòries, mètodes de Monte-Carlo, dinàmica molecular i mètodes basats en regles geomètriques combinades o no amb processos de minimització energètica.

En aquesta tesi s'han utilitzat principalment mètodes de dinàmica molecular. No obstant, en alguns casos concrets s'ha fet un anàlisi sistemàtic d'una part concreta de la molècula, i quan la dinàmica molecular no ha estat possible s'han utilitzat recerques aleatòries.

La *dinàmica molecular* consisteix en la simulació del moviment dels àtoms d'una molècula d'acord amb les lleis del moviment de Newton.¹³¹ Generalment, les forces que actuen sobre els àtoms es calculen fent servir camps de forces de mecànica molecular. Té la seva base en la hipòtesi ergòdica que afirma que "el promig temporal d'una propietat macroscòpica en un sistema és igual a la mitjana d'aquesta propietat en un sistema microscòpic (col·lectivitat canònica) en un instant de temps". Així doncs i per exemple, la distribució de valors en la rotació d'un enllaç és la mateixa tan si es consideren els valors de l'evolució en el temps d'un únic enllaç com si es consideren els valors de múltiples enllaços en un instant de temps.

5. Descriptors moleculars.

Els *descriptors moleculars* representen alguna propietat fisico-química de les molècules mitjançant un valor numèric. En l'àmbit de les relacions estructura-activitat s'utilitzen descriptors que caracteritzen els tipus d'interaccions més habituals en la unió substrat-receptor: hidrofòbic, estèric, electrostàtic i interaccions per enllaç d'hidrogen.¹³⁴⁻¹³⁶

Es poden diferenciar dos tipus de descriptors: Clàssics i 3D. Els primers descriuen les propietats moleculars a través d'un paràmetre monodimensional. Els segons descriuen les propietats mitjançant la projecció d'aquestes propietats en l'espai tridimensional que envolta les molècules. En aquest segon grup destaquen els *camps d'interacció molecular* (MIF).¹⁵⁶

En aquesta tesi s'han utilitzat exclusivament el camp d'interacció molecular GRID.

5.1. GRID.

Els camps d'interacció molecular GRID¹⁵⁷⁻¹⁶¹ proporcionen una eina computacional per determinar les zones d'interacció energèticament favorables de molècules conegudes. Es poden utilitzar tan per estudiar molècules individuals com fàrmacs, sistemes periòdics com membranes o cristalls, macromolècules, proteïnes, etcètera.

El procediment consisteix en posar la molècula o el sistema d'estudi (*target*) dins d'una caixa tridimensional en la qual es defineix una malla. En el primer punt de la malla es situa una molècula o entitat molecular, que s'anomena sonda (*probe*). Aquesta sonda pot ser una molècula com per exemple l'aigua, un fragment molecular com per exemple un residu

aminoàcid o quelcom sense representació física com per exemple la sonda hidrofòbica. En cada punt de la malla es calcula l'energia d'interacció entre la molècula d'estudi i la molècula o entitat molecular anomenada sonda. Les energies queden recollides en un vector (Figura 24).

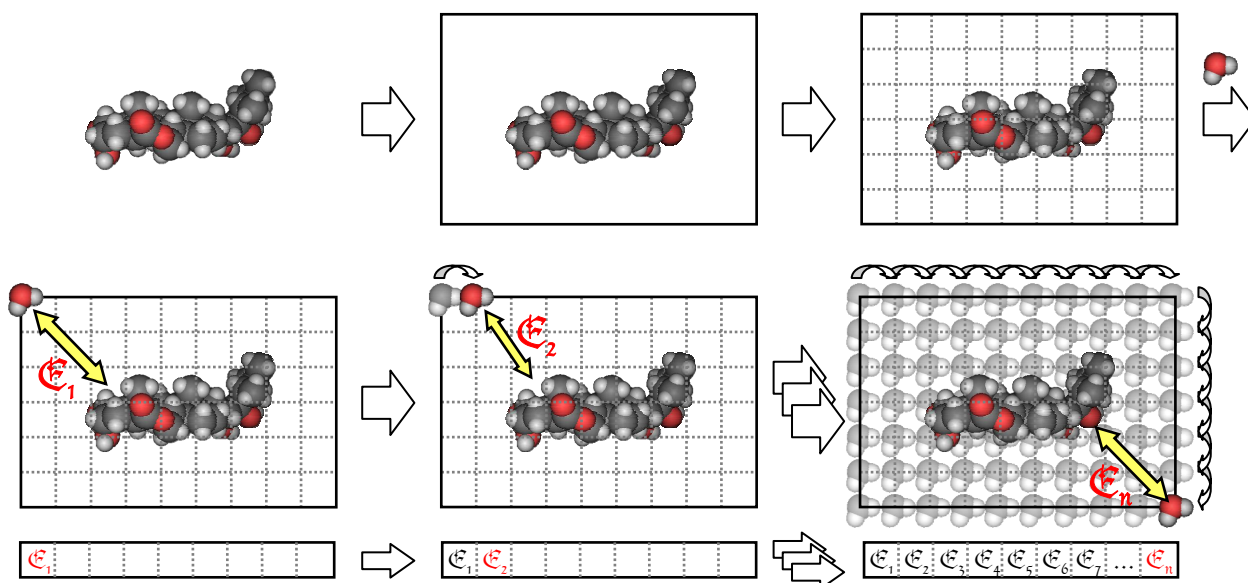


Figura 24: Esquema del procediment del càlcul dels camps molecular GRID fent servir la sonda d'aigua.

A diferència d'altres camps d'interacció molecular com per exemple CoMFA,¹⁶² GRID calcula l'energia d'interacció global. Es a dir, en cada punt es calcula l'energia d'interacció entre la molècula i la sonda com la suma dels diferents tipus d'interacció possibles: estèric, electrostàtic i enllaç d'hidrogen (Equació 1).

Equació 1:

$$E_{xyz} = \sum_{i=1}^{\text{àtoms}} E_{i-xyz}^{\text{estèric}} + \sum_{i=1}^{\text{àtoms}} E_{i-xyz}^{\text{electrostàtic}} + \sum_{i=1}^{\text{àtoms}} E_{i-xyz}^{\text{pont d'hidrogen}}$$

L'energia es pot visualitzar a un tall determinat (Figura 25). Per exemple, en la sonda d'aigua a -3 kcal/mol s'observen les interaccions electrostàtiques fortes i per enllaç d'hidrogen, mentre que a $-0,5$ kcal/mol s'observen principalment les interaccions estèriques.

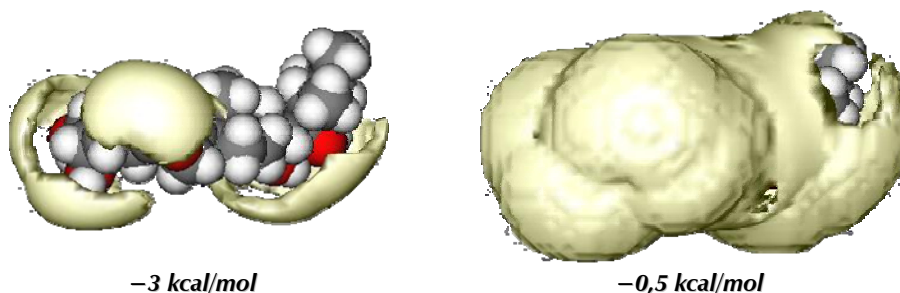


Figura 25: Visualització de l'energia d'interacció entre la brassinolida i l'aigua a -3 i $-0,5$ kcal/mol.

En aquesta tesi s'han fet servir les sondes d'aigua (molècula d'aigua) i metil (fragment molecular de metil). Tots els mapes de GRID que es mostraran corresponen a la sonda

d'aigua i estan representats a -3 kcal/mol per tal de poder discutir les interaccions electrostàtiques fortes i les interaccions per enllaç d'hidrogen.

5.2. Similitud molecular i Alineament.

El concepte de *similitud molecular* és en si mateix molt explícit, i tot i que se li han dedicat gran quantitat d'articles i llibres, en cap d'ells se l'ha arribat a definir de forma inequívoca.¹⁶³⁻¹⁶⁵ El motiu principal que argumenten és que la noció de similitud depèn considerablement del context en el qual s'utilitza. Similitud, dissimilitud o complementarietat són conceptes ambigus que necessiten ser definits i expressats matemàticament. En aquest sentit s'han descrit diversos índex de similitud entre els quals el de Carbó (C_{AB})¹¹³ i el de Hodgkin (H_{AB})¹⁶⁶ són dels més estesos. En aquests índexs l'oposat a la similitud (valor $+1$) no és la dissimilitud (valor 0) sinó la complementarietat (valor -1). Aquesta és respecte de la similitud el que un motlle al seu objecte.

L'índex de similitud de Carbó (C_{AB}) va ser originalment desenvolupat per comparar molècules d'acord amb la seva densitat electrònica. Posteriorment, s'ha generalitzat aquest índex de similitud per comparar qualsevol propietat estructural entre dos molècules (Equació 2a). Un altre índex de similitud que ha tingut molta acceptació és el de Hodgkin (H_{AB} , equació 2b). Mentre que l'índex de Carbó és sensible a la forma de la distribució de la propietat, el de Hodgkin reflecteix més la seva magnitud.

$$\text{equació 2a: } C_{AB} = \frac{\int P_A P_B}{\left(\int P_A^2\right)^{1/2} \left(\int P_B^2\right)^{1/2}}; \quad \text{equació 2b: } H_{AB} = \frac{2 \int P_A P_B}{\int P_A^2 + \int P_B^2}$$

El concepte d'*alineament molecular* està íntimament relacionat amb el de *similitud molecular* i és un dels punts crítics en els estudis de QSAR.

Per exemple, si es duplica una molècula (un BR), es gira 90° i es calcula els camp d'interacció de GRID amb la sonda d'aigua s'obté un vector on les zones d'interacció corresponents al carbonil de l'anell B no coincideixen (Figura 26). Llavors, es dona la paradoxa de que una molècula no és idèntica a si mateixa simplement pel fet de girar-la.

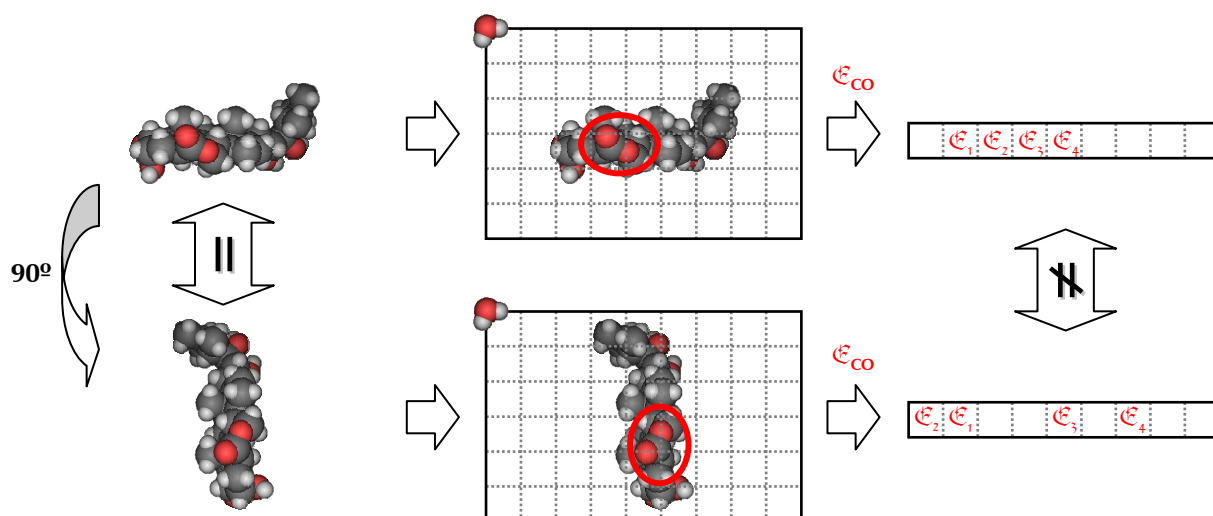


Figura 26: Exemple de la importància de l'alineament en els estudis de QSAR.

L'alineament molecular pretén corregir les diferències numèriques originades per una diferent distribució tridimensional de les molècules i es complica a mesura que augmenta la diferència estructural entre les molècules que es comparen. En aquests casos cal identificar les zones que són comunes a les dos molècules, la qual cosa no sempre és trivial, per utilitzar-les com a guia en el procés d'alineament.

El resultat de l'alineament depèn en gran mesura de la propietat triada per comparar i superposar les molècules entre si. Es poden emprar diverses propietats alhora de fer l'alineament:

- Esquelet comú: Les molècules s'orienten de forma que es minimitza la distància entre els àtoms equivalents que conformen un esquelet rígid, el qual és present en tots els compostos de la sèrie.
- Índex de similitud: Les molècules s'orienten de forma que es maximitza la superposició del potencial electrostàtic molecular (MEP) i/o el volum molecular definits a través d'uns índex de similitud.
- Camps moleculars: Les molècules s'orienten de forma que es maximitza la superposició d'un o diversos camps moleculars.

En l'optimització de la propietat escollida es pot considerar la molècula rígida o flexible. Si es considera rígida és necessari haver determinat amb anterioritat la conformació activa de la molècula. Si es considera flexible l'algorisme d'optimització ha de maximitzar l'índex de similitud alhora que minimitza també l'energia. Com tota optimització presenta el problema dels màxims locals de similitud.

En aquesta tesi s'ha utilitzat l'alineament a través d'un esquelet comú mitjançant la superposició dels anells C i D de l'esquelet esteroide (*Alineament CD*), i l'alineament de camps moleculars fent servir els camps de GRID amb la sonda d'aigua (*Alineament H₂O*). Aquest últim ha substituït l'alineament a través de la superposició de l'índex de similitud electrostàtic de Carbó (*Alineament MEP*) realitzat anteriorment en l'equip.

6. Tècniques estadístiques.

La *regressió* és la utilització de mètodes estadístics per crear un model matemàtic que descriu la variació d'una determinada resposta (variable dependent) com combinació dels efectes d'un conjunt de descriptors (variables independents).¹⁶⁷ La selecció de la tècnica estadística apropiada depèn de la naturalesa de la variable dependent, el nombre de descriptors i la correlació que pugui existir entre ells.

6.1. Regressió Lineal Múltiple (Multiple Linear Regression, MLR).

Tècnica estadística de regressió que descriu les variacions d'una resposta (variable dependent) com a combinació lineal de les variables independents. S'ajusta una equació lineal mitjançant mètodes de mínims quadrats. Les principals limitacions d'aquesta tècnica són: a) les variables han de ser linealment independents, no han d'estar correlacionades, i b) només es pot aplicar per un nombre reduït de variables, generalment una variable per cada 4 compostos a analitzar.

6.2. Anàlisi de Components Principals (Principal Component Analysis, PCA).

Es tracta d'una tècnica d'anàlisi multivariant d'un conjunt de dades mitjançant la identificació de combinacions lineals ortogonals entre si de les variables originals, per tal d'explicar la major part de la variabilitat de les dades. Aquesta tècnica "condensa" en uns pocs components principals la major part de la informació continguda en les variables originals. Aquesta visió simplificada del sistema permet descobrir pautes, tendències o agrupacions dels objectes que no són identificables mitjançant l'espai de les variables originals. A través de la interpretació dels components principals, es pot identificar les variables originals que tenen una major participació en la formació de les pautes o agrupacions.

Tot i que hi ha diversos algorismes, el més utilitzat consisteix en redefinir les variables inicials buscant, en primer lloc, una nova variable anomenada primer component principal (CP_1) que expliqui la màxima dispersió de les dades. Aquest component principal s'obté com una combinació lineal de les variables inicials. Posteriorment es poden definir més components principals que expliquin la màxima dispersió de les dades que encara queda per explicar. Els components principals són ortogonals entre si (Figura 27).

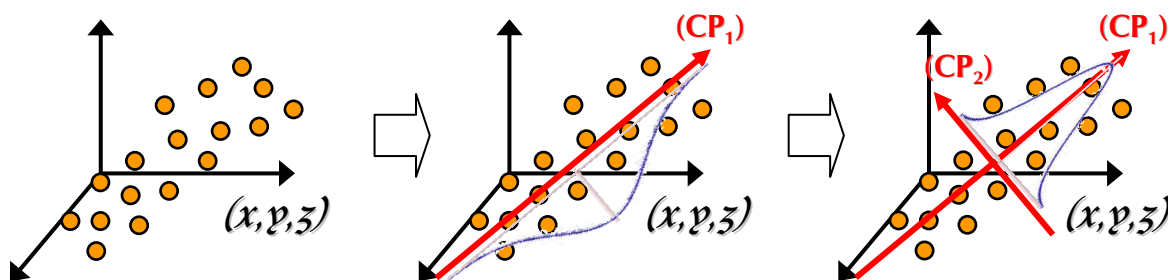


Figura 27: Exemple/Simplificació del funcionament del PCA. Les variables conformen els eixos de coordenades, sobre els quals es situen els objectes.

En el camp de 3D-QSAR s'utilitza per revelar les regions més importants d'un conjunt de molècules bioactives que han de ser tingudes en compte en el disseny d'anàlegs.¹⁶⁸

6.3. Regressió de Components Principals (Principal Component Regression, PCR).

Tècnica de regressió que consisteix en realitzar successivament un anàlisi de components principals (PCA) seguit d'una regressió lineal múltiple (MLR). Els components principals del PCA són utilitzats com a variables independents en la MLR. Això permet reduir el nombre de variables independents a uns pocs components principals, els quals en ser ortogonals es garanteix que són realment independents entre si.

6.4. Mínims Quadrats Parcial (Partial Least Squares, PLS).

Tècnica de regressió multivariant robusta que, igual que el PCA, redefineix les variables inicials buscant primer una nova variable anomenada variable latent (LV_1) que explica la màxima dispersió de les dades que alhora tingui relació amb les dades d'activitat (Figura 28).¹⁶⁹⁻¹⁷¹

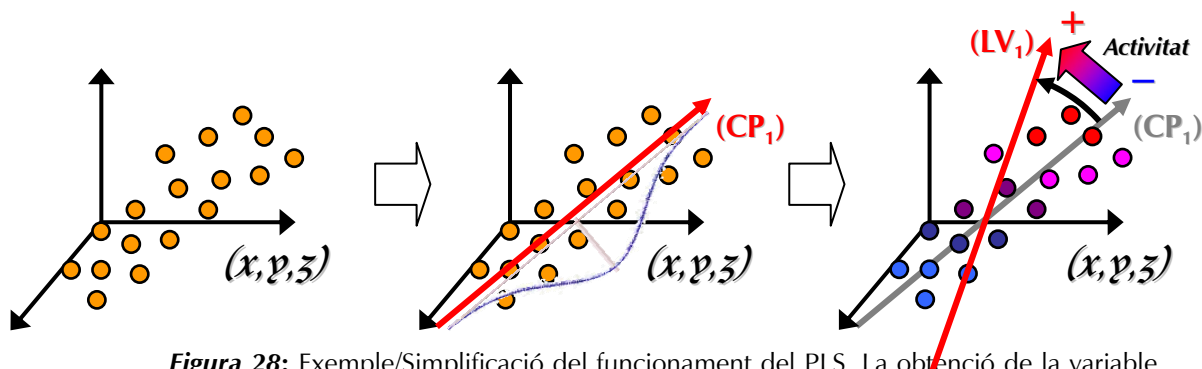


Figura 28: Exemple/Simplificació del funcionament del PLS. La obtenció de la variable latent ve guiada per les dades d'activitat.

6.5. Validació dels models de regressió.

La validació d'un model s'utilitza per determinar la seva capacitat tan de càlcul com de predicció. La capacitat de càlcul es mesura a través de la desviació estàndard de l'error de càlcul (*SDEC*, equació 3a) i del coeficient de correlació (r^2 , equació 3b). La capacitat de predicció es mesura a través de la desviació estàndard de l'error de predicció (*SDEP*, equació 4a) i del coeficient de predictibilitat (q^2 , equació 4b).

$$\begin{aligned} \text{equació 3a: } SDEC &= \sqrt{\frac{\sum (y_i^{\text{calc}} - y_i^{\text{real}})^2}{N}}; & \text{equació 3b: } r^2 &= 1 - \frac{\sum (y_i^{\text{calc}} - y_i^{\text{real}})^2}{\sum (y_i^{\text{calc}} - y_i^{\text{promig}})^2} \\ \text{equació 4a: } SDEP &= \sqrt{\frac{\sum (y_i^{\text{pred}} - y_i^{\text{real}})^2}{N}}; & \text{equació 4b: } q^2 &= 1 - \frac{\sum (y_i^{\text{pred}} - y_i^{\text{real}})^2}{\sum (y_i^{\text{pred}} - y_i^{\text{promig}})^2} \end{aligned}$$

Un dels objectius dels estudis estructura-activitat és predir l'activitat de nous anàlegs que encara no hagin estat sintetitzats. Per tant, la mesura en la capacitat de predicció d'un model és un bon criteri per avaluar-ne la seva qualitat.

Hi ha dos grans mètodes generals de validació:¹⁷²

- *Externa*: Utilitzen compostos diferents per construir el model i validar-lo. Abans de fer el model es diferencien dos grups de molècules. Un conjunt (*data set*) s'utilitza per construir el model de regressió. L'altre conjunt (*test set*) s'interpol·la en el model anterior i es contrasta l'activitat predita amb l'experimental.
- *Interna o validació creuada*: Utilitzen els mateixos compostos per construir el model i validar-lo.

En aquesta tesi i degut a limitacions en el nombre de BRs a estudiar, s'han utilitzat principalment mètodes de validació creuada. El més utilitzat és l'anomenat Leave One Out (LOO), tot i que en un cas concret s'ha fet servir el Leave Two Out (LTO) i la validació per grups. També s'ha assajat la possibilitat de reduir els models per fer servir mètodes de validació externa.

El mètode de validació interna Leave One Out (LOO) consisteix en construir una sèrie de models reduïts igual al nombre de molècules. En cadascun d'ells es deixa un compost fora del model reduït (d'aquí el nom). Posteriorment cada compost s'interpol·la en el model reduït del qual n'ha quedat fora i es contrasta l'activitat predita amb l'experimental. El mètode LTO és similar, però en cada model reduït es deixen dos compostos fora. El

mètode de validació per grups també és similar i és l'usuari qui defineix els grups que es queden fora de cada model reduït.

7. Activitat biològica.

7.1. Activitat a nivell general i a nivell molecular.

A nivell general l'activitat biològica d'un compost ha de ser considerada com l'efecte total que aquest compost provoca en l'organisme.¹⁷³ No obstant, l'estudi d'aquest efecte total és molt complex degut al gran nombre de processos involucrats i interrelacionats entre si. Per tant, es fa una simplificació i es considera només l'acció principal que el compost provoca en l'organisme.

A nivell molecular l'activitat biològica es correspon habitualment a la unió d'un compost de baix pes molecular (substrat) a un receptor macromolecular, generalment una proteïna (receptor). La unió substrat-receptor pot ser descrita com un procés d'associació i dissociació mitjançant un equilibri termodinàmic (Figura 29). En conseqüència l'activitat biològica s'expressa com l'afinitat entre les parts i la constant termodinàmica és una mesura de l'afinitat per substrat.



Figura 29: Esquema de la unió substrat-receptor i equació de l'equilibri termodinàmic. k_1 i k_{-1} són les constants d'associació i dissociació respectivament.

Per a que la unió entre un substrat i el seu receptor sigui productiva ha d'haver entre tots dos una complementarietat tan geomètrica com electrostàtica en el moment de la unió. Això dona lloc al concepte d'**especificitat per substrat**. En el centre d'unió existeix una quiralitat inherent deguda a que les proteïnes estan constituïdes per *L*-aminoàcids. Això fa que els centres d'unió siguin asimètrics i que la complementarietat electrostàtica no estigui garantida per estereoisòmers. La complementarietat geomètrica també ha d'existir i constitueix un requisit tan o més rigorós que l'anterior.¹⁷⁴

La corba de la resposta biològica en funció de la concentració de substrat (corba dosi-resposta) segueix una funció sigmoïdal. Al 50% de la resposta màxima la meitat dels receptors estaran lliures i l'altre meitat estaran ocupats. Això fa que la constant termodinàmica del procés sigui igual a la concentració de substrat on la resposta és el 50% de la resposta màxima i en conseqüència l'energia lliure de Gibbs associada a la unió substrat-receptor també està relacionada amb aquesta concentració (Figura 30). Aquesta concentració es coneix com $LD_{50\%}$, és una mesura de l'afinitat per substrat de les proteïnes, i és el paràmetre recomanat en els estudis de correlació estructura-activitat.¹⁷⁵

$$\Delta G_R^\circ = RT \log [K_D] = RT \log \left[\frac{k_1}{k_{-1}} \right] = RT \log \left[\frac{[S][R]}{[S \cdot R]} \right] \xrightarrow{50\% \text{ Rta}_{\text{màx}} \Rightarrow [R]=[S \cdot R]} \Delta G_R^\circ = RT \log [S_{50\%}]$$

Figura 30: Relació entre la energia lliure de Gibbs de la reacció i la concentració de substrat on la resposta és el 50% de la resposta màxima.

7.2. Tipus de bioassajos.

- Assajos “in vivo”.

Els assajos “in vivo” es realitzen amb el bioorganisme sencer i viu. En conseqüència en l'activitat no només es mesura la unió substrat-receptor sinó que també es tenen en compte altres factors com l'absorció del principi actiu, el transport i la seva distribució al llarg de l'organisme, el processos metabòlics i de degradació, etcètera. Tots aquests fenòmens tenen lloc en major o menor grau quan s'aplica el compost i són una caixa negra fins que no es tradueixen en un efecte macroscòpic observable.

Aquests assajos caracteritzen l'acció principal de l'activitat biològica global, però es desconeix l'efecte de cadascun dels processos implicats. El valor d'activitat pot continuar considerant-se com la concentració de substrat on la resposta és el 50% de la resposta màxima com a dada d'activitat ($LD_{50\%}$) ja que està relacionada amb l'energia lliure de Gibbs associada a tot el bioassaig (Figura 31).

$$\Delta G_{BA} = \Delta G_{Unió} + \Delta G_{Transp.} + \Delta G_{Metab.} + \dots = RT \log(K_{Unió} \cdot K_{Transp.} \cdot K_{Metab.} \cdot \dots) = RT \log(K_{BA})$$

Figura 31: Energia lliure de Gibbs del bioassaig (ΔG_{BA}).

- Assajos “in vitro”.

En els assajos “in vitro” cal diferenciar entre els que es realitzen amb una part d'un ésser viu (òrgan, teixit...) i els que disposen del receptor aïllat. En els primers la discussió és similar als assajos “in vivo”. En els segons, es pot estudiar específicament la interacció substrat-receptor determinant tant les constants cinètiques com les constants termodinàmiques. Aquests últims caracteritzen perfectament l'activitat a nivell molecular, però no l'activitat a nivell global.

7.3. Activitat i SAR.

En els estudis de relació estructura-activitat on no es coneix el receptor cal aplicar mètodes indirectes d'estudi. Aquests porten implícit la necessitat d'haver de considerar que en l'estructura de les molècules d'estudi hi ha inclosa tota la informació necessària per explicar el seu comportament i que totes les molècules s'uneixen a *un únic* receptor mitjançant els mateixos punts d'unió. Si l'activitat s'ha determinat “in vivo” és evident que no hi ha implicat un únic receptor.

En general, s'assumeix que l'únic efecte que incideix en l'activitat és la unió substrat-receptor i que la resta d'efectes (absorció, transport, metabolisme...) són similars per a tots els compostos. En el cas on no es pugui garantir l'anterior, es considera que el receptor és un receptor virtual del bioassaig que es descriu mitjançant la constant termodinàmica K_{BA} , la qual engloba tots els processos que tenen lloc.

En aquesta tesi, la situació és la descrita en últim lloc ja que l'activitat dels BRs s'ha determinat “in vivo”, es fan servir mètodes indirectes per correlacionar-la amb l'estructura i no està garantit que els efectes d'absorció, transport, metabolisme... siguin similars per a tots els compostos. L'exemple més clar és la castasterona, la qual és el precursor de la brassinolida i és activa “per se”.^{176,177}